Dysfunction of stress responsive systems in somatization

Tak, Lineke Maria

Dysfunction of stress responsive systems in somatization

Lineke Maria Tak

Cover: '*Fragmentation*', 2009, by Marie-Thérèse Hattink (www.mthattink.nl)
Lay-out: Martijn Harleman (www.bytegraag.nl)
Printed by: Ipskamp Drukkers BV, Enschede, the Netherlands
ISBN: 978-94-6070-010-1

The work described in this thesis was performed at the Interdisciplinary Center for Psychiatric Epidemiology of the University Medical Center Groningen, University of Groningen, the Netherlands and supported by grants from the Netherlands Organisation for Scientific Research (grant no. 016.056.064) and the Nicolaas Mulerius Fund. Publication of this thesis was supported by the Graduate School for Health Research SHARE.

RIJKSUNIVERSITEIT GRONINGEN

Dysfunction of stress responsive systems in somatization

Proefschrift

ter verkrijging van het doctoraat in de

Medische Wetenschappen

aan de Rijksuniversiteit Groningen

op gezag van de

Rector Magnificus, dr. F. Zwarts,

in het openbaar te verdedigen op

woensdag 20 januari 2010

om 14.45 uur

door

Lineke Maria Tak

geboren op 30 april 1983

te Apeldoorn

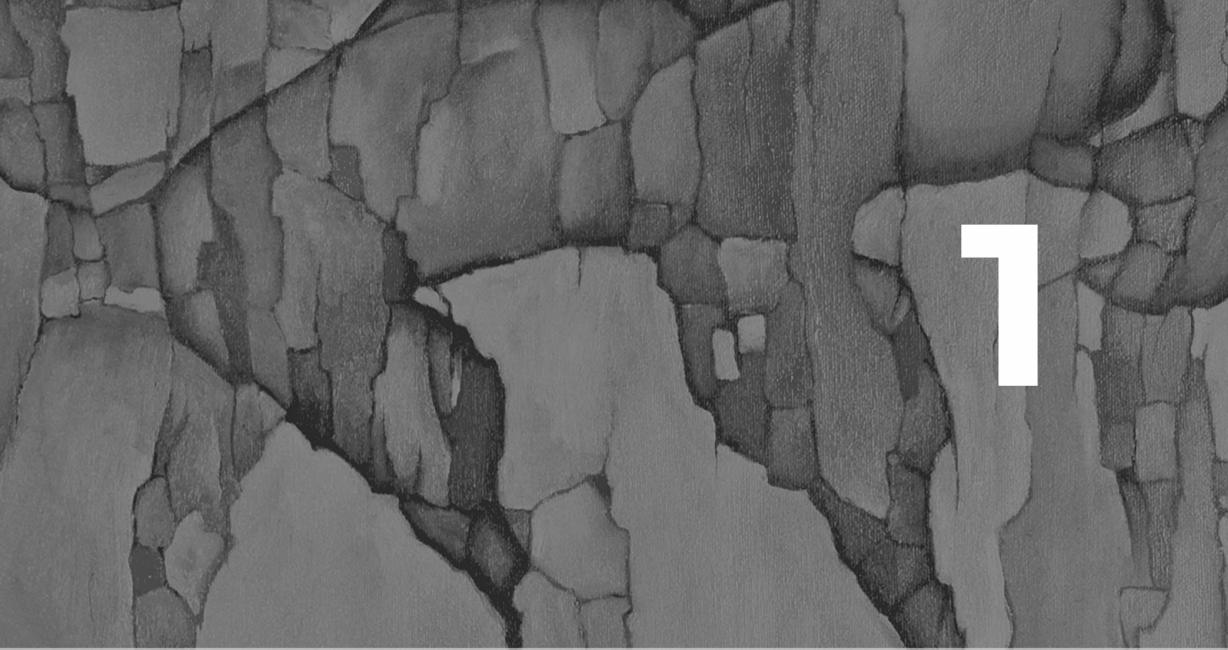
Promotores:	Prof. dr. J. Ormel Prof. dr. J.P.J. Slaets
Copromotor:	Dr. J.G.M. Rosmalen
Beoordelingscommissie:	Prof. dr. L.J.P. van Doornen Prof. dr. Ph. Spinhoven Prof. dr. R.O.B. Gans

It is devilish to suffer from a pain that is all but nameless.
Blessed are they who are stricken only with classifiable diseases!
Blessed are the poor, the sick, the crossed in love, for at least other people know what is the matter with them and will listen to their belly-achings with sympathy.
(In: 'Burmese Days', George Orwell, 1934)

Dr. Cameron: Dr. House doesn't like dealing with patients.
Dr. Foreman: Isn't treating patients why we became doctors?
Dr. House: No, treating illnesses is why we became doctors. Treating patients is what makes most doctors miserable.
(In: 'House, MD', TV series, Season 1, 2004)

Aan mijn ouders

CONTENTS



1

CHAPTER 1

General introduction

LM Tak and JGM Rosmalen

Revised and condensed version of chapter published in:
Psychoneuroendocrinology Research Trends, Hauppage, New York: Nova Science Publishers, 2007;463-95

CONCEPT OF SOMATIZATION: A CONTINUUM WITH MANY FACES

The purpose of this chapter is to define and describe somatization and its related constructs, to illustrate the clear need for a better understanding of somatization and to provide background information to explain the rationale behind studying the role of stress responsive system dysfunction in relation to somatization. The chapter ends with the main research questions and content of this thesis.

Functional somatic symptoms

Symptoms are in the heart of clinical medicine as a consult with a physician typically starts with the question "what are your complaints?" Physicians structure their differential diagnosis on the constellation of symptoms that a patient presents. A symptom is defined as an aversively perceived internal state (van Wijk & Kolk 1997), such as abdominal pain, fatigue, nausea, dizziness or muscle weakness. Normal physiological fluctuations and pathophysiological processes may trigger interoceptors that generate information about the state and function of organ systems. Most of those somatosensory stimuli remain unconscious but some become conscious and are perceived as a sensation or as a symptom.

Experience of symptoms in the general population is an extremely common phenomenon. In a population survey, 96% reported that they had experienced at least one symptom during the preceding month (Ihlebaek et al. 2002). Symptoms may or may not be accompanied by objective signs of underlying disease such as findings derived from physical examination, blood tests, or diagnostic imaging. When an individual experiences symptoms for which no known organic pathology (i.e., a deviation from the normal condition in a part, organ, or system of the body characterized by an identifiable group of signs or symptoms) can be found, these symptoms are referred to as functional somatic symptoms (FSS). Also when organic pathology is present, symptoms may have a functional component when they are in excess of what would be expected based on objective findings (Kroenke & Rosmalen 2006).

Although the prevalence of FSS in the general population is high, only a minority of people seek medical attention for them (Kroenke & Price 1993, Hiller et al. 2006). Still, FSS are the most common single category of symptoms in primary care and form a substantial part of complaints presented in medical specialties (Kroenke & Mangelsdorff 1989, Katon & Walker 1998, Nimnuan et al. 2001a). Precise estimates are dependent on the population under study and definition of FSS. Typically, a third to half of the presented symptoms remains unexplained in primary care and population based studies, and of this group of FSS, 25% are chronic or recurrent (Kroenke & Rosmalen 2006).

Although a large part of symptoms presented to doctors is functional somatic, it appears that doctors find it difficult to satisfactorily explain this to patients and

reassure that their symptoms are as real as symptoms with a known organic cause (Hartz et al. 2000). Alongside and partly due to those insufficient explanations of doctors, patients experiencing FSS tend to equate the explanation that no organic pathology is found for their symptoms with 'not real', 'simulated', or 'imaginary'. Psychosocial and somatic factors are incorrectly believed to be mutually exclusive and symptoms with a known organic pathology are experienced as more legitimate than FSS. Patients often resist clarifications that there is no organic pathology causing their symptoms, or that stress is possibly involved. This is particularly the case when a functional explanation is introduced for the first time after all somatic examinations have failed to provide results that confirm organic disease (Bensing & Verhaak 2006).

Sometimes no clear-cur stressors are present, and moreover, it is difficult to understand why stress would lead to abdominal pain, a feeling of a lump in throat, or muscle weakness. Under the assumption that there must be organic pathology underlying their symptoms, patient might clamp to the diagnostic process and consult many health care providers for the same problem. Consequently, patients with FSS disproportionately use health care resources (Barsky et al. 2005). No uniform intervention or treatment protocol for FSS across primary care, somatic medical specialties and psychiatry is available. FSS cause a large burden in patients experiencing them. FSS are often as disabling in daily life as somatic symptoms with a known organic cause (Kisely et al. 1997). Patient with FSS have a poorer self-rated health than people presenting the same symptoms with a known pathological basis (Frostholm et al. 2007).

In aggregate, FSS form a highly prevalent and clinically important and costly health care issue, which frustrates physicians and dissatisfies patients. Clearly, it is essential to better understand the etiology of FSS.

Functional somatic disorders

FSS tend to occur together and result in functional somatic disorders (FSD), syndromes of related complaints with no known organic pathology. Without logical nomenclature based on underlying pathology, FSD are often described by their lead symptoms or implied cause. The main three FSD are chronic fatigue syndrome (CFS), fibromyalgia (FM), and irritable bowel syndrome (IBS). Other examples of FSD include whiplash syndrome, functional dyspepsia, temporomandibular dysfunction, tension-type headache, multiple chemical sensitivity, non-cardiac chest pain, hyperventilation syndrome, reactive hypoglycemia, low back pain, and atypical facial pain. History seems to repeat itself, as new disorders with clusters of FSS continue to be defined. Some diagnoses have only recently been acknowledged as FSD, such as interstitial cystitis and chronic non-bacterial prostatitis, however, this has not been generally accepted yet (Clemens 2008). The introduction of a new FSD is inevitably followed by a search to causal factors, such as exposure to micro-organisms or

chemical agents (Iversen et al. 2007). These steps are similar to the search for an isolated causal factor in CFS, FM, and IBS, that started more than a decade ago and is still proceeding in some research groups.

It is a matter of debate whether all FSD are in fact manifestations of one single disorder, as pursued by the 'lumpers', or that FSD cannot be narrowed in a single entity, as argued by the 'splitters'. FSD overlap in case definition, reported symptoms, and in non-symptom characteristics such as gender, prognosis, and response to treatment (Wessely et al. 1999, Barsky & Borus 1999, Aaron & Buchwald 2001, Fink et al. 2007, Henningsen et al. 2007), supporting the view of the lumpers. The splitters postulate that despite of all the commonalities, the differences between the FSD cannot be ignored. They consider the finding that specific infections and premorbid levels of stress differentially precipitated IBS and CFS supportive for their view (Moss-Morris & Spence 2006). Those two views are not mutually exclusive. Shared factors might underlie general susceptibility for development of any FSD, whereas FSD-specific factors might shape their final manifestation (Aggarwal et al. 2006, Kato et al. 2009). Despite this discussion, it seems generally accepted that FSD have at least a common core, allowing researchers to include different FSD when discussing determinants of the tendency to experience FSS in general.

Somatoform disorders

FSS can also be diagnosed on the psychiatric multi-axial diagnostic system of the Diagnostic and Statistical Manual, Fourth Edition (DSM-IV) (American Psychiatric Association 1994). In this classification system, FSS are classified in the category of somatoform disorders, together with conversion disorder, hypochondriasis, and body dysmorphic disorder. Although the latter three diagnoses are categorized under somatoform disorders in DSM-IV, those diagnoses are dissimilar from the FSS count diagnoses and their relationship is debatable (Mayou et al. 2005), and are therefore not considered in this thesis.

In the DSM-IV, FSS are defined as 'physical complaints not fully explained by a general medical condition, by the direct effects of a substance, or by another mental disorder'. In all somatoform disorders, FSS must cause clinically significant distress or impairment in social, occupational, or other areas of functioning. In contrast to factitious disorder and malingering, the somatic symptoms are not intentional or feigned.

Diagnosis of somatization disorder requires a history of many FSS beginning before the age of 30 years that occur over a period of several years. Each of the following criteria must have been met, with individual FSS occurring at any time point during the course: four pain symptoms, two gastro-intestinal symptoms, one sexual symptom, and one pseudoneurological symptom. Undifferentiated somatoform disorder is characterized by one or more FSS lasting at least six

months with the number of FSS below the threshold for a diagnosis of somatization disorder. The diagnosis of pain disorder is made in patients when pain has existed for at least six month in at least one anatomical site. Psychological factors are judged to play a role in the onset, severity, exacerbation, or maintenance of the pain. DSM-IV does not inform users about the kind of psychological factors and who has to judge whether they play a role. Somatoform disorder not otherwise specified (NOS) codes disorders with FSS that do not meet the criteria for any of the other somatoform disorders. Somatoform disorder NOS can be diagnosed when somatoform symptoms are present but criteria for another somatoform disorder are not met, such as in case of FSS of recent onset or short duration.

While FSS are common, somatization disorder is a very rare diagnosis with an estimated prevalence of 0.03 – 0.84% in the general population (Creed & Barsky 2004). Prevalences of undifferentiated somatoform disorder and somatoform disorder NOS have not been surveyed in the general population, but almost 30% of primary care consulters meet the diagnostic criteria (Fink et al. 1999).

In response to the broadly interpretable definitions of undifferentiated somatization disorder and somatoform disorder NOS in the DSM-IV, researchers have been introducing several new definitions that require a less-restrictive than eight, but variable number of FSS. Examples include Abridged Somatization Disorder (Escobar et al. 1998), Somatization Symptom Index (Escobar et al. 1989), Bodily Distress Disorder (Fink et al. 2007), and Multisomatoform Disorder (Kroenke et al. 1997). Whilst uniformity in classification is lacking, these proposals provide the implicit notion that counting the exact number of FSS seems redundant and somatization is better seen as a continuum. Empirical studies that support such a dimensional model have been published (Katon et al. 1991, Liu et al. 1997, Hiller et al. 2006).

Somatization

Historically, somatization is defined as the tendency to communicate emotional distress in physical or bodily terms (Lipowski 1988). This view assumes that emotional distress is involved and unhelpfully underscores dualistic thinking. Somatization is better described as the processes involved in experiencing somatic symptoms not or not conclusively explained by known organic pathology (Lipowski 1988) (i.e., FSS) along a continuum from relatively mild symptoms to severe and disabling disorders (Hiller et al. 2006).

As the etiology of somatization is not understood, it seems logical to maintain neutrality in its nomenclature. No generally accepted definition for somatic symptoms not or not conclusively explained by organic pathology exists. The long list of descriptions, ‘functional somatic symptoms’, ‘medically unexplained symptoms’, ‘symptoms that are difficult to objectify’, ‘inexplicable health

problems', 'somatization symptoms', 'symptoms without substrate', 'somatoform symptoms', 'subjective health complaints', 'psychogenic symptoms', 'vague symptoms', and 'psychosomatic symptoms' illustrates the ongoing discussion about semantics but also about etiology. Definitions expressing a psychological connotation, such as 'psychogenic symptoms' and 'psychosomatic symptoms', are preferred by neither researchers or patients. 'Medically unexplained symptoms' is regarded honest and neutral by some researchers, however, many patients dislike this term (Stone et al. 2002). Moreover, the term 'medically unexplained symptoms' contradicts with a lot of models - more or less supported by research - that provide explanations of symptoms in terms of interacting biological, psychological, and social processes (Brown 2007). As earlier applied, we use the term 'functional somatic symptoms' (FSS), since this definition side steps the psychological versus physical dichotomy, is etiologically neutral, and is preferred by patients (Stone et al. 2002).

Alike FSS, disorders with clusters of FSS have parallel classification systems and definitions that are confusing. FSD in somatic medicine, such as CFS, FM, and IBS, suggest specific organic disease and thus highlight specific bodily patterns. Somatoform disorders in psychiatry highlight psychological processes and only count the number of FSS (Mayou & Farmer 2002). Patients may be classified by both a FSD and a somatoform disorder for the same set of FSS (Mayou et al. 2005). In the psychiatric classification system for instance, a patient with musculoskeletal FSS can be diagnosed with a somatoform disorder on axis I as well as with fibromyalgia on axis III. There is a clear need to harmonize the isolated classification systems; nevertheless, there is little agreement on which label is the most appropriate.

In this thesis, somatization is defined as a broad concept. When using the term somatization, we refer to people in the general population presenting with relatively mild FSS, as well as to patients with FSD, and, at the very end of the spectrum, to somatization disorder according to DSM-IV.

Etiology of somatization

Little is known about elementary processes in the physiology of symptom experience. Visceral signals, afferent sensory pathways processing those signals to the central nervous system, and higher-order cognitive processes that are involved before a symptom is reported are largely unexplored (Barsky 2001). Several mechanisms have been proposed to explain somatization. Briefly, those mechanisms can be distinguished in vulnerability factors, triggering factors, and perpetuating factors (Deary et al. 2007).

Vulnerability factors make a person susceptible for somatization and include genetic predisposition (Gillespie et al. 2000), early life experiences with parental illness (Hotopf et al. 1999), early adverse life events such as sexual abuse or

childhood maltreatment (Salmon & Calderbank 1996), and neuroticism or other stress sensitive personality traits (Rosmalen et al. 2007, Turk et al. 2008).

Triggering factors are catalysts for somatization in already vulnerable persons, and include sensitization (Ursin 1997), acute psychological stressors, infections or trauma (Moss-Morris & Spence 2006, Roelofs & Spinhoven 2007), chronic psychosocial stress (van Houdenhove & Egle 2004), and attention processes, such as health anxiety, depression, self-focused attention, or lack of distraction (Pennebaker 1982, Lowe et al. 2008).

Perpetuating factors maintain the somatization process and include a predominant somatic attribution style, dysfunctional health cognitions, and iatrogenic somatic fixation due to diagnostic tests and referrals (Rief et al. 2004). Behavioral processes such as symptom-led activity pattern, dysfunctional coping strategies, and social reinforcement also contribute to perpetuation of somatization (Vercoulen et al. 1997, Kirmayer & Looper 2006).

It has also been hypothesized that FSS are nothing more than somatic representations of anxiety and depressive disorders. Indeed, high rates of co-morbidity exist between FSS on the one hand and anxiety and depressive disorders on the other hand (Haug et al. 2004, Lowe et al. 2008). In a study in primary care, however, only 26% of the patients with a somatoform disorder - mainly undifferentiated somatoform disorder - had an anxiety or depressive disorder (Waal de et al. 2004). Furthermore, a meta-analysis indicates that FSD are related to, but certainly not fully dependent on anxiety and depressive disorders (Henningsen et al. 2003).

FSS are not fully explained by any single mechanism and a multifactorial approach is needed. Recent research and theory in this area show complex interactions between physiological, psychological, and social factors in the development and perpetuation of FSS. Many models incorporating those factors have been advanced and all have different features. Importantly, they all share the idea of different levels in the pathway of symptom perception: disturbances in generation, attention, interpretation, and behavioral processes may all account to varying extents in different persons.

Relatively little attention has been paid to underlying biological mechanisms. When appreciating research concerning biological mechanisms, two directions of research can be identified. The first direction concerns brain imaging studies. Structural and functional brain imaging both point to central nervous system alterations in somatization, which are extensively reviewed (Wood 2005). The second direction concerns studies that are indicative of alterations in responsiveness of important bodily stress responsive systems. This thesis will focus on the latter direction of research.

We will start the next section by outlining the rationale to investigate stress responsive system dysfunction in somatization, followed by a summary of available evidence for this association.

STRESS RESPONSIVE SYSTEM DYSFUNCTION IN SOMATIZATION: AN INCONCLUSIVE AREA OF PSYCHOSOMATIC RESEARCH

The rationale to investigate stress responsive systems in somatization

The first studies investigating stress responsive system function in somatization were published about twenty years ago. A reason to investigate an etiological link between the stress responsive systems and somatization emerges from the potential of those systems to increase symptom experience and perception (Sharpe & Bass 1992). Often, the idea of stress responsive dysfunction as a mediator between psychosocial stress and somatization has been advanced, under the assumptions that psychosocial stress is associated with somatization and that among the long-term effects of psychosocial stress are chronic under- or over activity of the stress responsive systems.

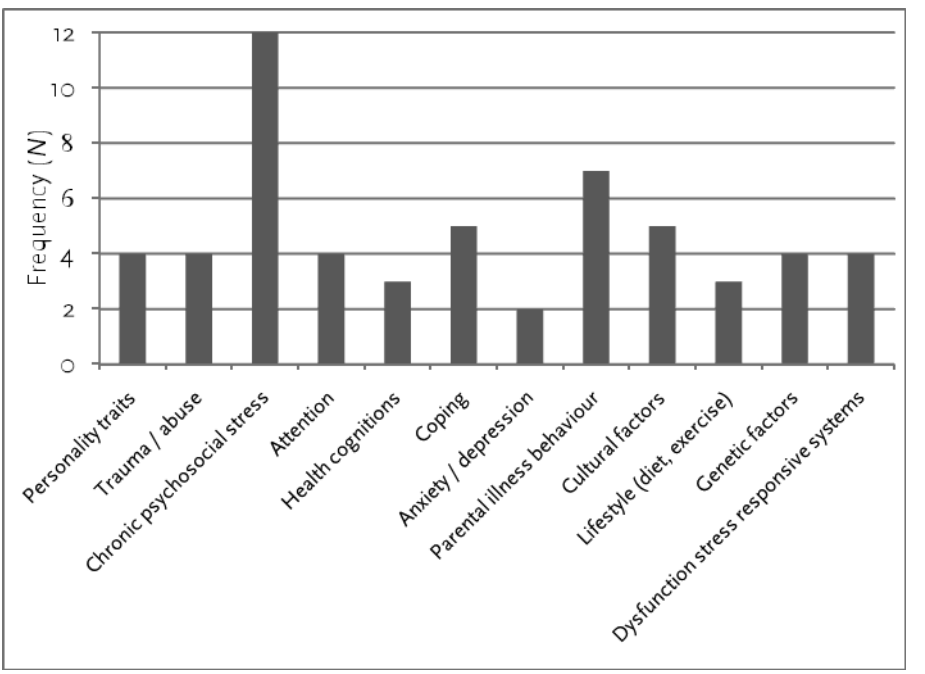
We begin by outlining the rationale for this theory. For each stress responsive system, we will start with a brief description of the most important features, followed by a discussion of potential consequences of dysfunction for symptom experience and available evidence for such dysfunction in somatization. Then, we will discuss limitations of the current evidence and explore alternative pathways.

Psychosocial stress and somatization

To start with the first assumption, an important theory about the etiology of somatization is the chronic stress hypothesis. The chronic stress hypothesis refers to the generally existing idea that stress has part in the development and experience of FSS and FSD. Stress has been defined as environmental demands that individuals appraise as heavily taxing or exceeding their coping abilities (Lazarus & Folkman 1984). This definition clearly indicates that stress comprises two components: the presence of a stressor (i.e., demand from an individual's environment) and the subjective experience of stress due to the presence of this stressor (i.e., the appraisal that the stressor taxes the coping abilities). The nature of stressors may be predominantly physiological or predominantly psychosocial. Psychosocial stress is the overall name for stressors that have primarily a psychological or social nature. Usually, a distinction is made between major life events, chronic difficulties, and daily hassles. Examples of major life events are death of a spouse or a divorce (Brugha et al. 1985). Examples of chronic difficulties are work overload or permanently living with insufficient financial resources (Hendriks et al. 1990). Examples of daily hassles are losing things and minor family concerns (Kanner et al. 1981). It seems widely accepted that

psychosocial stress impacts on development of symptoms and disorders that are not explained by organic pathology. In the general population, FSS are often referred to as 'stress-related symptoms' by laypersons (www.helpguide.org, 2009). Professionals, such as primary care physicians (Wileman et al. 2002) and psychiatrists (see Figure 1) also think that psychosocial stress is the most important etiological factor in FSS. Indeed, virtually all models on the etiology of somatization include forms of psychosocial stress. The question whether psychosocial stress precedes FSD, such as CFS, FM, and IBS, has been answered with a "fairly unequivocal yes" in a narrative review (Deary et al. 2007). Stress sensitive personality traits, premorbid psychosocial stress, also if happened several decades ago, and acute psychosocial stress predict development of CFS (Heim et al. 2006, Kato et al. 2006, Prins et al. 2006), FM (Wolfe & Hawley 1998, Gupta & Silman 2004, van Houdenhove et al. 2005), and IBS (Mayer 2000, Blanchard & Scharff 2002, Folks 2004). Life events also precede different FSD surveyed in the general population (Aggarwal et al. 2006).

Figure 1. *Etiology of somatization according to psychiatrists.*



Before the start of a symposium on concepts and mechanisms of somatization, 18 psychiatrists and psychiatrists in training were asked to write down which etiological factors they considered important in development of functional somatic symptoms and disorders. Only factors that were reported by at least two psychiatrists are shown.

Stress responsive system dysfunction in somatization

The second assumption is that chronic stress is able to induce alterations in stress responsive system activity. When stress exceeds a certain threshold, it evokes a generalized stress system response. The aim of this response is to restore homeostasis (Chrousos & Gold 1992). The autonomic nervous system, the hypothalamic-pituitary-adrenal axis, and the immune system all serve to protect the body from stress and can therefore be seen as bodily stress responsive systems (McEwen 1998b). When the load of stressors in an individual is too large, when stress responsive systems are chronically addressed, or when the capacity of the stress responsive systems to adjust is diminished, dysfunction of stress responsive systems may develop (Chrousos & Gold 1992). Stress responses are generated by a network of integrative brain structures involving the paraventricular nucleus of the hypothalamus and the amygdala. These structures receive input from visceral and somatic afferents and from cortical structures. This integrative network provides output to the pituitary and to the pontomedullary nuclei, structures that mediate the autonomic and endocrine output to the body. Repeated or prolonged exposure to stress may result in changes in this central stress circuitry. Consequently, these changes may result in stress responsive system alterations.

Autonomic nervous system (ANS)

The ANS is responsible for rapid stress responses, since it reacts within seconds after stimulation. The ANS controls bodily functions such as thermoregulation, breathing, and circulation. It helps maintain homeostasis and coordinates responses to external stimuli. The ANS can be divided in two divisions: the sympathetic and the parasympathetic nervous system. The sympathetic nervous system is frequently referred to as the 'stress' or 'fight or flight' system, as it has a stimulatory effect on bodily systems and organs that are responsible for quick sensory activity and movement. The sympathetic influence is mediated through the postganglionic release of noradrenalin. Roughly said, the parasympathetic nervous system antagonizes the sympathetic nervous system, since it has 'rest and digest' activity. Stimulation of the parasympathetic nervous system results in bradycardia and relaxes many bodily systems and organs through the postganglionic release of acetylcholine. The ANS does not function on its own, but is both anatomically and functionally linked to other parts of the nervous system.

ANS activity is influenced by chronic, repetitive, and acute psychosocial stress (Sloan et al. 1994, Dishman et al. 2000, Schommer et al. 2003). ANS alterations as a result of psychosocial stress would most likely consist of decreased activity of the 'resting' parasympathetic nervous system and increased activity of the 'stress' sympathetic nervous system. As the ANS innervates several organs, sensation and misinterpretation of peripheral physiological arousal may result in experience of FSS (Sharpe & Bass 1992, Rief & Barsky 2005). Examples of FSS that may result

from misinterpretation of autonomic physiological arousal are functional chest pain in case of increased heart rate, functional abdominal pain in case of decreased gastro-intestinal peristalsis, and functional musculoskeletal pain in case of increased muscle tension.

A widely used proxy for ANS function is heart rate variability (HRV), reflecting interbeat interval fluctuations in heart rate (HR). In an attempt to study sympathetic nervous system activity, studies have often reported power in the low frequency band (HRV-LF), defined at 0.04 - 0.15 Hz. However, the physiological basis is not well understood and HRV-LF certainly does not simply reflect sympathetic activity, although frequently reported as such. Therefore, we here consider only studies reporting on resting PNS activity in the high frequency band (HRV-HF), defined at 0.15 - 0.40 Hz and also referred to as cardiac vagal activity, as this measurement has a clear physiological basis (Berntson et al. 1997).

A narrative review of more than thirty studies suggests that decreased cardiac vagal activity could be associated with presence of the main FSD (i.e., CFS, FM, and IBS), however, findings are not fully consistent (Tak & Rosmalen 2007). When HRV is measured under resting conditions, if there are any significant differences, they point consistently in the direction of decreased cardiac vagal activity in FSD compared to healthy controls. Nevertheless, studies often find no baseline differences between patients with FSD and healthy controls. Although reliability of HRV measurements performed during challenges is generally poorer than when measured at rest (Sandercock et al. 2005b), several strategies to challenge the ANS have been used to provoke HRV responses, such as measurements during standing, tilt table testing, deep breathing, isometric exercise, treadmill walking, thermal stimuli, and mental stress. Also, specific challenges or procedures, such as eating a meal or rectal balloon distention, have been performed in IBS studies. In about half of the studies, no differences after challenge tests are apparent. When significant differences were found compared to healthy controls, cardiac vagal activity was always lower or responsiveness was decreased in FM and CFS, and mostly in IBS. Only one study in IBS found increased cardiac vagal activity in cases compared to controls (Tousignant-Laflamme et al. 2006). Another study specifically looking at IBS symptom groups found that those groups were characterized by different physiological responses to food intake (i.e., an increase in cardiac vagal activity in constipation-predominant IBS and a decrease in cardiac vagal activity in diarrhea-predominant IBS) (Elsenbruch & Orr 2001b). No studies on ANS function in somatization disorder, pain disorder, undifferentiated somatoform disorder, or somatoform disorder NOS have been performed. Furthermore, no studies on the association between ANS function and the extent of FSS in individuals from the general population are available.

Hypothalamic-pituitary-adrenal axis (HPA axis)
In response to several afferent stimuli, the hypothalamus releases corticotrophin-releasing hormone (CRH) and arginin vasopressin (AVP). CRH and AVP synergistically induce release of adrenocorticotropic hormone (ACTH) from the pituitary. ACTH is secreted with a diurnal rhythm superimposed upon the pulses of CRH. The lowest serum ACTH concentrations occur in the early night, whereas the highest ACTH concentrations occur between 0400h a.m. and 0600h a.m. In addition to this diurnal rhythm, ACTH responds to a wide variety of stimuli. ACTH controls the release of cortisol from the adrenal cortex. The diurnal curve of total and free plasma cortisol includes 7 to 13 pulses of cortisol secretion per day. Half of the total daily cortisol is secreted within the major burst before dawn. Cortisol causes negative feedback of its own secretion at several levels, including hippocampus, hypothalamus, and pituitary. The biologically active component of cortisol is the unbound or free fraction and the main biological effects are widespread, including increase of glucose levels and blood pressure, fat metabolism, protein catabolism, breakdown of muscle mass, lowering the activity of the immune system, and altering mood. These biological effects depend on many factors other than blood plasma concentrations, such as sensitivity of glucocorticoid receptors and the presence of other molecules. It is important to realize that the HPA axis does not function on its own, but interacts with many other bodily systems (Niewoehner & Bantle 1998, Genuth 1998).

There are several hypothetical pathways how altered concentrations of corticotrophin-releasing hormone (CRH), adrenocorticotropic hormone (ACTH), and cortisol can be involved in somatization. Firstly, CRH not only modulates the endocrine response, but also influences pain perception. Although acute stress is known to produce analgesia, chronic stress may have the opposite effect, a process mediated by CRH (Clauw & Chrousos 1997, Lariviere & Melzack 2000). Low cortisol concentration may cause widespread pain and fatigue (Heim et al. 2000a, Fries et al. 2005).

In this thesis, the focus is specifically on cortisol, of which secretory patterns can provide a partial window into the activity of the HPA axis. Cortisol can be assessed in blood plasma, saliva, or urine. Measurement can take place at a single time point, but it is generally preferred to collect cortisol several times a day, or during several days (Nicolson 2007). Although there is no gold standard, assessment of cortisol in saliva seems to be the preferred method as it has proven to be a valid and reliable reflection of the respective unbound concentration in blood and has advantages such as stress-free sampling and laboratory independence (Kirschbaum & Hellhammer 1994). Twenty-four hour urinary free cortisol (24-h UFC) excretion is considered a practical index of integrated plasma free cortisol, not influenced by the time of measurement and interpersonal differences in circadian rhythm (Levine et al. 2007).

Although more than sixty HPA axis studies in FSD have been published, they do not exclusively support one separate kind of dysfunction. Narrative reviews conclude that baseline findings on cortisol levels in CFS, FM, and IBS subjects compared to healthy controls are inconsistent: mild hypocortisolism seems the most frequent finding, but normal or increased cortisol levels have also been reported (Mayer et al. 2001, Geenen et al. 2002, Cleare 2003, Tak & Rosmalen 2007).

Dynamic challenges of the HPA axis in FSD have been undertaken, to detect more subtle disturbances and to determine the level of dysfunction in the HPA axis that cause reported cortisol alterations FSD in rest. The HPA axis can be activated by several stressors and findings have been presented on different pharmacological challenges, including the insulin tolerance test, CRH stimulation test, vasopressin infusion, standard ACTH stimulation test (250 ug), low dose ACTH stimulation test (1 ug), and glucagon stimulation test. In addition, cortisol measurements after administration of medication, such as d-fenfluramine, naloxone, and buspirone have been performed. To challenge the HPA axis non-pharmacologically, measurements after procedures to elicit physical or psychological stress have been used, such as auditory stress, food intake stress, exercise stress, thermal stress, rectal balloon distention, sigmoidoscopy, mental tasks, or public speaking. In narrative reviews on HPA axis activity after stress tests, divergent results were reported, without clear evidence for any specific change to the HPA axis. When studies find differences, mostly an impaired or blunted response of the HPA axis has been shown (Cleare 2003, Tanriverdi et al. 2007). One of the theories regarding the underlying cause of hypocortisolism in FSD is that of enhanced negative feedback on the hypothalamus or pituitary. To test this theory, the dexamethasone suppression test is most often used, in which enhanced negative feedback in FSD compared to controls indeed is the main finding (Tak & Rosmalen 2007).

No studies regarding HPA axis activity in somatization disorder, pain disorder, undifferentiated somatoform disorder, or somatoform disorder NOS exist. However, two studies including patients with less strict diagnoses than somatization disorder have been performed. The first study observed no alterations in a range of cortisol measurements in patient with at least eight FSS compared to healthy controls (Rief & Auer 2000). The second study even observed higher basal salivary cortisol concentrations in patients having at least four FSS (when male) or at least six FSS (when female) compared to healthy controls (Rief et al. 1998). In a study making 24-h real-life ambulatory recordings of salivary cortisol in patients with different FSD (such as hyperventilation syndrome and CFS, but also other conditions such as burnout syndrome), no differences compared to healthy controls were found (Houtveen & van Doornen 2007b). In a community study of 41 adults, plasma cortisol levels measured after the dexamethasone suppression test did not predict self-perceived fatigue. No

other studies on HPA axis function and FSS in the general population are available.

Immune system

A third important stress responsive system to consider in the etiology of FSS is the immune system. Unlike the ANS and the HPA axis, the immune system is not primarily known as a stress responsive system. The immune system is a complex system known for its protection against pathogens and can be divided into innate and adaptive immunity (Male 2001). Innate immunity is immediate but non-specific, whereas adaptive immunity is specific and recognizes a pathogen. Immune responses are produced primarily by leucocytes and soluble mediators of which there are several different types. Phagocytes, such as monocytes and macrophages, internalize micro-organisms and are a component of the innate immunity. Lymphocytes, such as B cells and T cells, have more specialized functions and are central to adaptive immunity. Cytokines are proteins that enable immune cells to communicate and play an integral role in the initiation, perpetuation and subsequent down regulation of the immune response. Cytokines can be divided into pro-inflammatory cytokines such as interleukin (IL)-1, tumor necrosis factor (TNF)-α alpha, and IL-6, and anti-inflammatory cytokines like IL-4 and IL-10 (Male 2001). Production of C-reactive protein (CRP), an acute phase protein in the immune reaction, is induced by pro-inflammatory cytokines. CRP levels may rise rapidly and as much as 1000-fold after an acute inflammatory stimulus (Black et al. 2004). It has been documented that the immune system can become triggered by non-pathogenic, more generic signals, such as psychological stress (Kiecolt-Glaser et al. 2002). A meta-analysis showed that acute stressors are associated with potentially adaptive upregulation of the immune response in a similar pattern for men and women across the life, whereas chronic stressors were associated negative effects on almost all functional measures of the immune system, especially in older adults (Segerstrom & Miller 2004). This meta-analysis did only include few studies on cytokines and none on CRP and chronic stress. In a population-based study of older adults, higher levels of chronic stress were associated with higher concentrations of IL-6 and CRP (Ranjit et al. 2007).

Different pathways may be etiologically involved in the link between the immune system and somatization. Firstly, the immune system may be related to FSS and FSD via sickness behavior, a constellation of non-specific symptoms induced by pro-inflammatory cytokines, such as IL-1, IL-6, and TNF-α (Dantzer 2001, Wieseler-Frank et al. 2005). These non-specific sickness behavior symptoms, including fatigue, weakness, malaise, hyperalgesia, and increased focus on own body, are core characteristics of persons having FSS and FSD (Dantzer 2005). Secondly, immune activation may be a marker for activity of the ANS and HPA axis, as both systems closely interact with the immune system (McEwen et al. 1997, Araujo et al. 2006).

A lot of other immunological abnormalities (e.g. T cell quantity and function, B cell quantity and function) have been identified in several FSD; however, those abnormalities are rarely replicated. This is illustrated by a critical review of immunological abnormalities in CFS. No consistent pattern of immunological abnormalities has been identified, and additionally, the high quality studies where the ones that demonstrated that cytokine levels in patients were not different from those in controls (Lyall et al. 2003). Furthermore, a study combining patients with different FSD did not find differences in pro- and anti-inflammatory cytokines between somatizers and healthy controls (Houtveen et al. 2007a). In a less restrictive diagnosis of somatization syndrome, another study found lower levels of pro-inflammatory cytokine IL-6 in somatization patients compared to healthy controls (Rief et al. 2001), an unexpected direction based on theory. In sum, although elevated cytokine levels may induce sickness behavior, they are not unambiguously being found in somatization. Maybe these conflicting findings are due to the fact that alterations in central behavior of cytokine, if present, are not necessarily measurable in the peripheral blood (Dantzer 2001). Therefore, other immunological biomarkers that reveal stress responsiveness of the immune system are important to consider.

CRP might be a more integrated and accurate peripheral marker for innate immune system activation. Although historically long been considered as clinically irrelevant, minor elevations of CRP (3 - 10 mg/L) have been reported to be associated with psychosocial stress (McDade et al. 2006). Ultrasensitive assays can detect CRP in this subclinical range as high-sensitive CRP (hs-CRP). Thus, elevated hs-CRP does not directly causes generation or sensation of FSS but rather is a biomarker of immune activation. Of associations of circulating concentrations of CRP and pro-inflammatory cytokines in subjects of the general population without active infection, those between IL-6 and CRP have been most frequently investigated and are generally strongest, with reported significant correlation coefficients ranging from 0.24 to 0.50 (Ridker et al. 2000, Cesari et al. 2003, Piche et al. 2005, Berrahmoune et al. 2007, Stewart et al. 2008, Milaneschi et al. 2009). In the study with the strongest correlation between IL-6 and CRP, which included 991 subjects aged 65 years and older, correlations of CRP with other investigated cytokines were lower: significant correlations of 0.09 and 0.19 were found for IL-1β and IL-18 respectively, and an non-significant correlation of 0.06 was found for TNF-α (Milaneschi et al. 2009). Indeed, the few available findings addressing hs-CRP appear more consistent. Patients with FSD have higher levels of hs-CRP compared to healthy controls (Buchwald et al. 1997, Spence et al. 2008), although this difference is not always statistically significant, such as in IBS (Schoepfer et al. 2008). No data about immune function or hs-CRP in persons experiencing FSS in the general population have been published.

Methodological problems
Summarizing the available results on the association between the stress responsive systems and somatization, one encounters heterogeneous results. Nevertheless, some specific directions can be distilled from research findings on the ANS and the HPA axis. If associated at all, somatization seems quite consistently characterized by lower cardiac vagal activity and hypocortisolism. However, research focusing on stress responsive system dysfunction and somatization is mainly conducted in small samples derived from clinical settings, which are not representative for all individuals with FSS. The wide variability in assessment of different parameters, particularly when measuring HPA axis or immune system activity, may also contribute to the mixed findings. Furthermore, little attention has been paid to the role of potential confounders, such as body mass index, smoking, physical activity, sleeping problems, psychiatric co-morbidity, concurrent psychosocial stress, and medication use, or potential moderators, such as gender and age (Tak & Rosmalen 2007).

Causality and alternative pathways
It remains unclear whether stress responsive system dysfunction, if present, is causally related to somatization. The temporal relationship between stress responsive system dysfunction and somatization is not well established, because previous studies have generally been cross-sectional. Several pathways should be considered; an overview of these four pathways is shown in Figure 2.

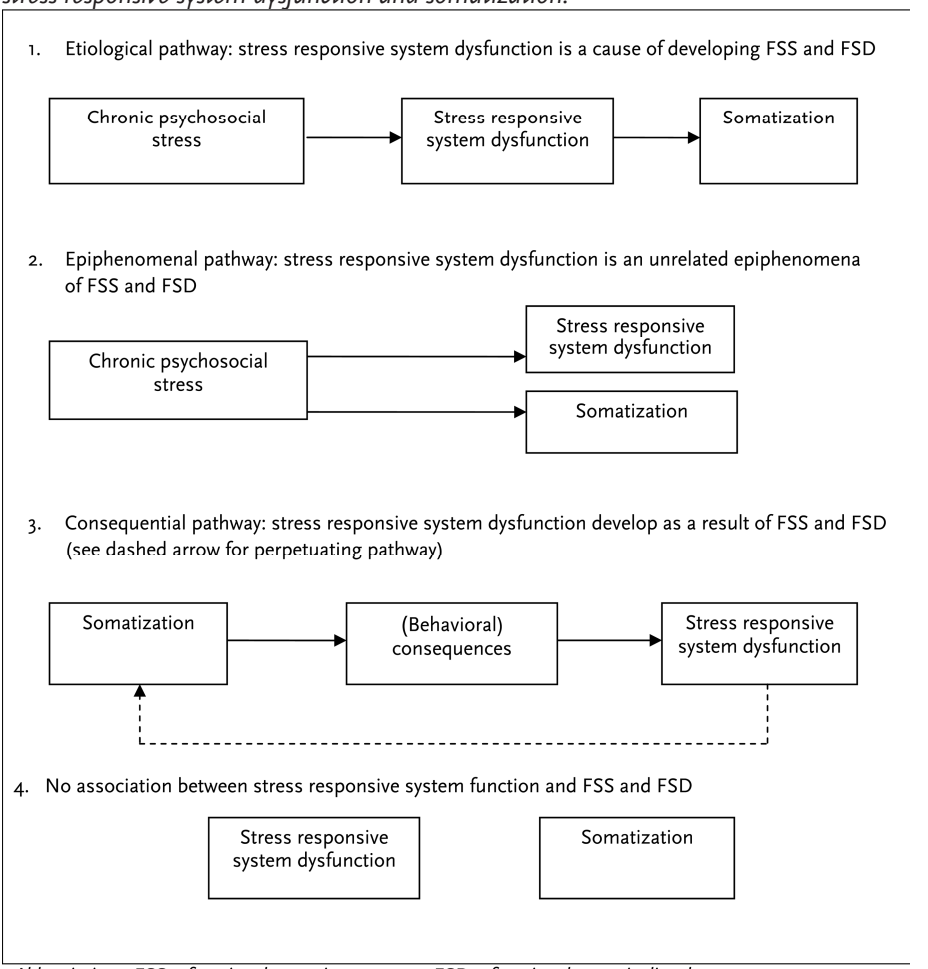
Firstly, alterations in stress responsive systems could induce FSS: the etiological pathway. In this chapter, we have introduced the hypothesis of stress responsive dysfunction as a causal risk factor in the etiology of somatization. Although this etiological pathway is biologically plausible, preceding and etiologically important stress responsive system dysfunction may also arise from other factors than psychosocial stress, such as genetic predisposition.

Somatization and stress responsive system dysfunction could also share the same etiological determinants, such as stress sensitive personality or psychiatric co-morbidity, and develop parallel in the epiphenomenal pathway. In this pathway, stress responsive system dysfunction is not causally involved in somatization but only a correlate.

Thirdly, FSS and FSD could induce stress responsive system alterations, for example, due to symptom experience or lifestyle alterations: the consequential or concomitant pathway. In this pathway, stress responsive system alterations may not cause symptom experience at all and, whereas it may also be a perpetuating factor of symptom experience, and constitute a combination of the consequential and etiological pathway (i.e., perpetuating pathway).

Finally, it may be possible that there is no association between stress responsive system dysfunction and FSS and FSD at all. Previous positive findings may be artifacts of methodological shortcomings and publication bias.

Figure 2. *Four simplified representations of theories about the relationship between stress responsive system dysfunction and somatization.*



Abbreviations: FSS = functional somatic symptoms, FSD = functional somatic disorders

We are primarily interested in stress responsive system dysfunction in the etiology of somatization. Accordingly, it is of importance to assess whether stress responsive system dysfunction satisfies the requirements of being a causal risk factor. To assess this, a decision tree for classifying the association between a certain factor and an outcome has been proposed (Kraemer et al. 1997). A risk factor is a measurable characterization of each subject in a specified population

that precedes the outcome of interest. A characterization that satisfies all requirements for a risk factor except for precedence is a correlate of the outcome. This is a crucial point, because concomitants or consequences of outcomes are likely more highly correlated with the outcomes than are risk factors. A risk factor that can change spontaneously within a subject (such as age or body mass index) or that can be changed with an intervention is a variable risk factor. A risk factor that cannot be changed (such as genetic predisposition or gender) is a fixed risk marker. The term causal risk factor can only be used when a variable risk factor is manipulable and change the risk of the outcome when it is manipulated (see for a schematic overview Figure 3).

AIMS OF THIS THESIS

Hypotheses

The overall aim of the studies described in this thesis is to test the presence of dysfunction of stress responsive systems in the etiology of somatization. We hypothesize that dysfunctions in stress responsive systems, as characterized by hypofunction of the HPA axis, altered ANS function in the direction of decreased cardiac vagal activity, and increased immune system activation, are shared causal risk factors in the etiology of FSS and the main FSD.

Investigating whether there is a causal role for dysfunction of stress responsive systems in the etiology of somatization is only possible in large cohort studies, a type of design that is uncommon in this research field. Therefore, we started with extracting as much as possible available information from cross-sectional case-control studies by using meta-analysis. This approach is particularly valuable to investigate whether stress responsive system dysfunction is a shared or FSD-specific correlate and which other factors influence their association. Furthermore, we used data from a large longitudinal population-based cohort study to evaluate whether stress responsive system dysfunction is a prospective risk factor for FSS.

Outline

Following this introductory chapter, *Chapter 2* focuses on the general principles of performing a meta-analysis, with an emphasis on its specific merits and pitfalls for observational, psychosomatic research. *Chapter 3* is a systematic review of methodological quality and meta-analysis of ANS function in FSD. *Chapter 4* presents a meta-analysis and meta-regression on HPA axis activity in FSD. Both chapters aim to test whether stress responsive system dysfunction is a shared or FSD-specific factor. *Chapter 5* tests the assumption that somatization may be better considered as a continuum instead of requiring FSS in several bodily clusters. Next, we report on the association between ANS function (*Chapter 6*),

HPA axis activity (*Chapter 7*), and immune function (*Chapter 8*) and FSS using prospective measurements in a population-based cohort (Figure 4). *Chapter 9* presents an exploratory study on the association between psychosocial stress and stress responsive system function. Finally, in *Chapter 10*, the main findings will be summarized and integrated with as central question: what is the evidence for a pivotal role of dysfunction of stress responsive systems in somatization?

Figure 3. *The process of elucidation of risk factor status for a factor in a population for a particular outcome.*

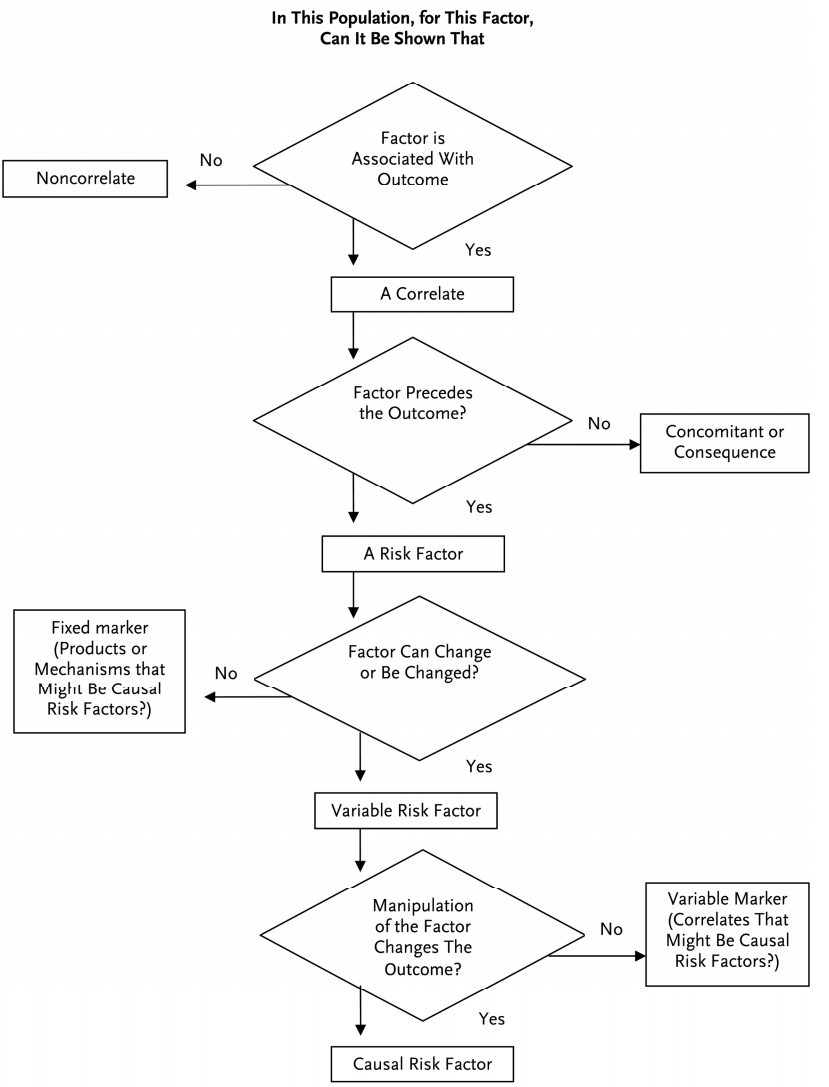
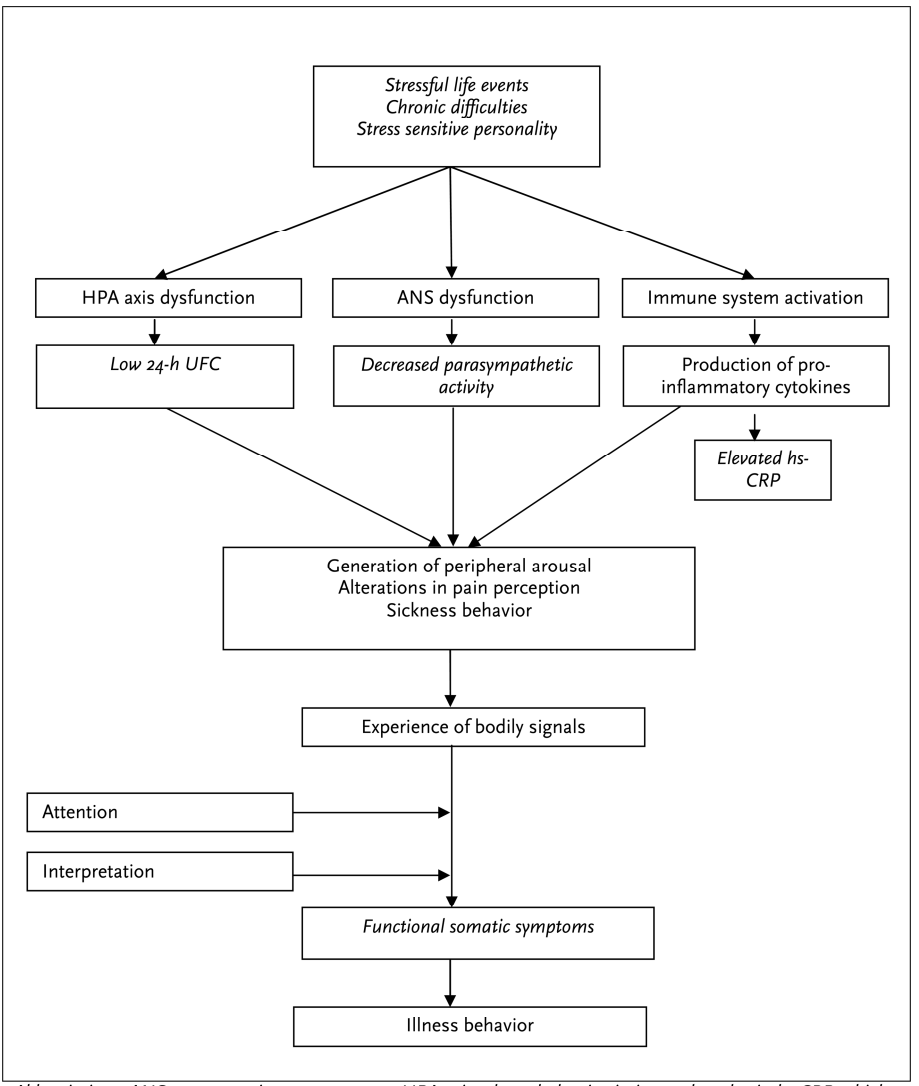
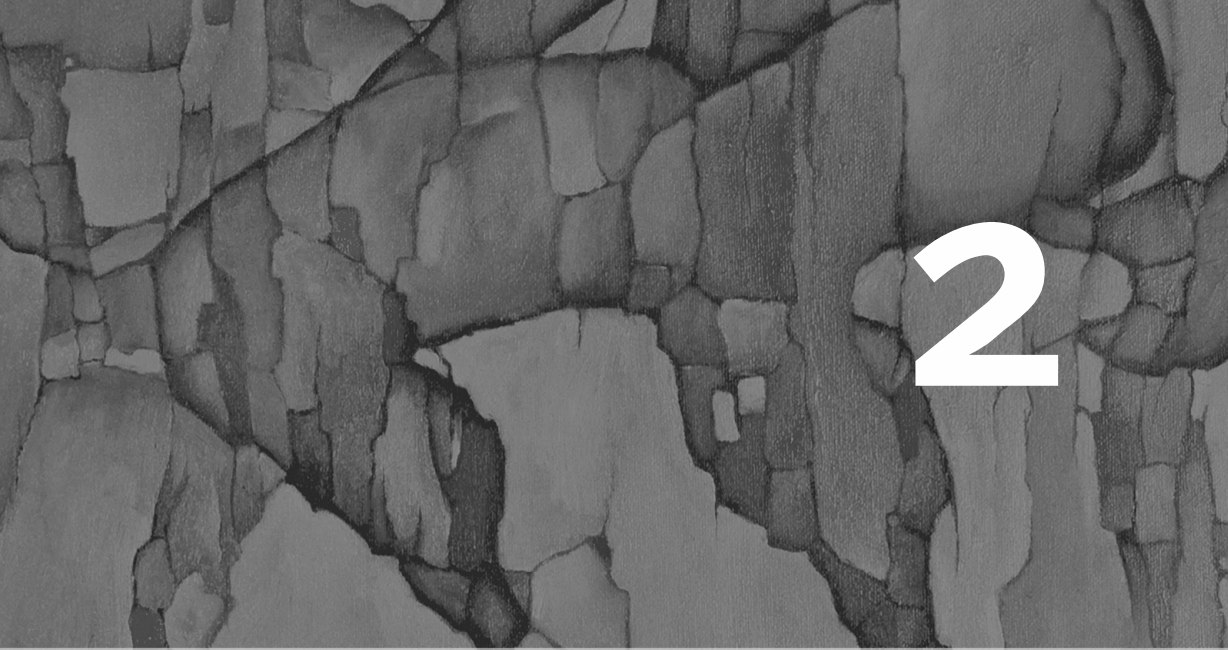


Figure 4. *Stress responsive system dysfunction and somatization. Variables in italic represent variables we have measured in the population-based studies in this thesis.*



Abbreviations: ANS = autonomic nervous system, HPA axis = hypothalamic-pituitary-adrenal axis, hs-CRP = high-sensitive C-reactive protein, 24-h UFC = 24-hour urinary free cortisol



2

CHAPTER 2

More than the sum of its parts: Meta-analysis and its potential to discover sources of heterogeneity in psychosomatic medicine

LM Tak, A Meijer, A Manoharan, P de Jonge, and JGM Rosmalen

Psychosomatic Medicine, in press

ABSTRACT

Meta-analyses may contribute to more reliable knowledge about the existence of certain relations in the area of psychosomatic research. Surprisingly, the increasing popularity of meta-analysis is not reflected in the number of meta-analyses on observational studies published in Psychosomatic Medicine. This may be due to the specific difficulties that apply to meta-analyses of observational research. The aim of this paper is to provide a non-technical overview of the principles of meta-analysis applied to observational research. We will highlight general principles of meta-analysis and discuss the major threats to its validity, with an emphasis on its specific merits and pitfalls for psychosomatic research using several examples. We conclude that meta-analysis is a relatively simple technique, leaving little reason for not routinely applying it when performing a systematic review. An adequately conducted meta-analysis may not only provide a summary estimate of a certain association, but has additional value in discovering relevant confounders, mediators, and moderators as well as identifying areas of research that require more attention.

INTRODUCTION

Meta-analysis is a statistical procedure that integrates several studies concerning a certain research question to reach a more secure conclusion. State-of-the-art meta-analyses have the potential to provide a more objective appraisal of the evidence than traditional narrative reviews. They can reveal that repeated results in the same direction across several studies, even if not one is significant, can be much more powerful evidence than a single significant result from an individual study (Rosenthal & DiMatteo 2001). In addition, meta-analyses can provide insight into why different studies have found different results. Furthermore, whereas adequately powered randomized controlled trials have a relatively high positive predictive value of reflecting the true relationship, this value drops for research findings of observational studies (Ioannidis 2005). Meta-analysis of reported associations in observational studies may raise the positive predictive value of the true relationship. Finally, meta-analysis may identify areas of research that need more investigation, for example based on results of subgroup or sensitivity analyses (Egger et al. 1997d, Garg et al. 2008).

Meta-analysis is a particularly useful procedure in psychosomatic research. Many studies in this field meet characteristics that contribute to the risk of non-replication, such as having small sample sizes, retrieving small effect sizes, testing a large number of relationships without clear rationale, and having large flexibility in designs, definitions, outcomes, and analytical methods (Ioannidis 2005, Freedland et al. 2009). In spite of this, the sharply increased number of published meta-analyses in general medicine (Egger & Smith 1997d, Sutton & Higgins 2008) is not reflected in Psychosomatic Medicine, where the average number of meta-analyses published is stable at around two per year (Table 1). Since 2000, 15 out of 16 meta-analyses published in Psychosomatic Medicine were purely based on observational research (and the other one was a combination of observational and intervention research (Cho et al. 2005)), whereas this was the case for only one fourth of the 127 meta-analyses published in this time period in the Journal of the American Medical Association, as an example of a general medical journal. In the light of certain difficulties that are associated with meta-analyses of observational studies, most importantly the risk to produce very precise but equally spurious results (Egger et al. 1998a), the relative low number of meta-analyses in Psychosomatic Medicine may not be surprising.

Since the last review on meta-analysis in Psychosomatic Medicine almost two decades ago (Rosenthal 1991), new techniques, procedures, and recommendations have become available (Sutton & Higgins 2008). of the general principles of meta-analysis to readers with a basic knowledge of statistics. The current paper provides an update in the form of a non-technical overview.

Table 1. *Overview of meta-analyses published in Psychosomatic Medicine since 2000.*

First author	Year	Association	*N* of primary studies	Effect size measure	Fixed or random effects	Homo- or heterogeneity	Moderator analyses	Meta-regression	Quality assessed	Publication bias
Howren	2009	Depression and inflammation	9-61	*d*	Random	Heterogeneity	5-10	No	-	Funnel plot, Fail safe *N*
Chida	2008	Positive wellbeing and mortality	19-21	HR	Random	Heterogeneity	5-10	No	+	Funnel plot, Egger's test, Fail safe *N*
Chida	2008	Psychosocial factors and atopy	9-34	*r*	Random	Heterogeneity / Homogeneity	5-10	No	+	Begg's test
Steffen	2006	Acculturation and blood pressure	114-124	*d*	Random	Heterogeneity	5-10	No	+	Funnel plot, Fail safe *N*
Carter	2006	Cesarean section and postpartum depression	8	OR	Not stated	Heterogeneity	0	No	+	Not assessed
Cho	2005	Placebo response in chronic fatigue syndrome	29	Proportion	Random	Heterogeneity	<5	Yes	+	Not assessed
Barth	2004	Depression in coronary heart disease and mortality	4-7	OR, HR	Random	Homogeneity / Heterogeneity	<5	No	-	Funnel plot
Van Melle	2004	Post-MI depression and cardiovascular prognosis	6-9	OR	Fixed / Random	Homogeneity / Heterogeneity	<5	No	-	Funnel plot, Egger's test
Dickens	2003	Depression and pain perception	2-6	*d*	Fixed	Homogeneity	0	No	-	Funnel plot
Sundin	2003	Major life events and stress by Horowitz Event Scale	66	n.a.	n.a.	n.a.	n.a.	Yes	-	Not assessed
Wulsin	2003	Depressive symptoms and coronary heart disease	10	OR	Random	Heterogeneity	0	No	+	Funnel plot, Begg's test, File drawer *N*
Henningsen	2003	Depression and anxiety in functional syndromes	3-45	*d*	Fixed / Random	Homogeneity / Heterogeneity	<5	No	-	Fail safe *N*
Dickens	2002	Rheumatoid arthritis and depression	9-12	*r*	Not stated	Heterogeneity	5-10	No	+	Fail safe *N*, File drawer *N*
Rutledge	2002	Psychological factors and hypertension	15	*r*	Random	Homogeneity	5-10	No	+	Fail safe *N*
De Groot	2001	Depression and complications of diabetes	22	*r*	Random	Heterogeneity	<5	No	-	Fail safe *N*
Buckley	2001	Cardiovascular measures in PTSD	4-28	*d*	Not stated	Homogeneity	<5	No	-	Not assessed

Abbreviations: HR = hazard ratio, MI = myocardial infarction, n.a. = not applicable, OR = odds ratio, PTSD = post traumatic stress disorder.

We will highlight the general principles of meta-analysis and discuss the major threats to its validity, with an emphasis on its specific merits and pitfalls for observational, psychosomatic research. Our aim is that after reading our paper, researchers should be able to critically interpret meta-analyses performed by others or are encouraged to perform a meta-analysis themselves.

INCLUSION OF STUDIES

Searching the literature

In searching literature for a meta-analysis, the goal is to include as many of the existing relevant studies as possible in a reproducible manner. Meta-analyses on psychosomatic research should use at least the literature databases Medline and Embase. In addition, other, more subject-specific databases can be searched, depending on the research question (e.g. PsycINFO, CENTRAL, CINAHL, ISI Science and Social Science Citation Index, Cochrane Library). Each database has specific search possibilities and most databases provide tutorials. Databases often use key-words, but free text should also be searched (Counsell 1997, Higgins & Green 2008). The search should be conducted without language restrictions to reduce the risk of language bias (Egger et al. 1997c, Juni et al. 2002). Because searching the literature is almost a specialty in itself and errors in search strategies are common (Sampson & McGowan 2006), additional consultation of a librarian may be worth considering.

The probability that research findings are published is influenced by the nature and the direction of the results. Significant research findings are overrepresented, whereas results conflicting with the prevailing beliefs about the association are underrepresented (Hardy & Thompson 1998). Searching for unpublished studies is thus important in order to achieve a representative sample of the work available in the research area under study, but it requires considerable effort. Unpublished findings may, for example, be revealed by asking relevant research groups for any unpublished results or checking dissertation databases (e.g. Dissertation Abstracts Online, ProQuest Dissertations and Theses).

Selecting relevant studies

Articles are selected for inclusion based on a predesigned protocol containing inclusion criteria specifying the type of subjects, exposure, outcomes, and type of study (Meade & Richardson 1997, Counsell 1997, Higgins & Green 2008). This is preferably done by two independent reviewers, since they select on average 9% more studies than one (Edwards et al. 2002). One particular problem in the selection process is the fact that several articles with different first authors may report on the same study, or on partly overlapping data. This problem may especially occur in observational studies that gather information on a large

number of variables over a relatively long period of time, resulting in more than one publication on a single study. Just like the search strategy, the selection process should be reported in detail.

Methodological quality

Critical appraisal of the methodological quality of primary studies is an essential feature of meta-analysis. Good methodological quality can be defined as having a design that minimizes bias in the estimation of the association under study. Critical appraisal checklists or scales ('tools') can be used as a threshold for inclusion of studies, or preferably, the meta-analysis can be repeated excluding low quality studies to assess whether results would change (see: sensitivity analysis). Although there is a plethora of tools for assessing quality of intervention trials, consensus on the ideal tool for assessing methodological quality of observational studies is currently not available (Sanderson et al. 2007). Major domains that should be incorporated in every observational studies quality tool are selection of participants, measurement of dependent variables, and control for confounding.

The type of tool used to assess quality can dramatically influence the interpretation of meta-analyses (Juni et al. 1999). In order to develop a valid tool, experts in the field could be consulted and development of the tool should be clearly stated. Reliability of the tool can be assessed by using at least two independent raters to score the individual papers and interrater agreement statistics should be reported. Researchers should be aware that using a quality tool inevitably introduces subjectivity, such as the decisions on which items to include and on the scoring rules for each quality item. When developing a quality tool, general items for quality of reporting can be used, such as consensus guidelines on reporting of randomized trials, CONSORT (Altman 1996); diagnostic tests, STARD (Bossuyt et al. 2003); or observational studies in epidemiology, STROBE (von Elm et al. 2007). Recommendations for adequate reporting of case-control studies in the psychiatric setting have also been made (Lee et al. 2007), which are largely applicable to observational psychosomatic research. Additionally, researchers can include specific items that are pivotal for good quality studies in their field. Assessing sources of bias is a crucial but equally complex function of a quality tool, since distinguishing quality of reporting and quality of the actual study design is often not possible. Notwithstanding some degree of uncertainty about the validity of comparing study quality, quality tools specifically designed for a meta-analysis of a certain topic under study may additionally serve as a guideline for conducting high quality future research.

An example of a quality tool for meta-analysis of observational studies in the psychosomatic area is one developed for studies on cardiac vagal activity in functional somatic syndromes. Experts in the field were asked to review this quality tool that includes items such as whether the functional somatic syndrome

has been reliably assessed, whether the measurement of cardiac vagal activity has been reported in appropriate units, and whether specific covariates such as age, gender, body mass index, depression, and medication use have been assessed or adjusted for (Tak et al. 2009a). In this meta-analysis, it could not be proven that study quality accounted for the mixed findings. It was advised, however, that future research adhering to the proposed quality criteria may provide a more definite answer on the question whether lower cardiac vagal activity is involved in the etiology of functional somatic syndromes. The assessment of methodological quality should be considered as a routine procedure in meta-analysis.

PERFORMING THE META-ANALYSIS

Effect size per study

The basic information needed for a meta-analysis is the effect size per study, which is the measure of the magnitude (strength) of the association between two variables. Information needed to calculate this effect size consists of a summary measure and a measure of its precision (standard error or 95% confidence interval). Widely used summary measures are the correlation coefficient, odds ratio (OR), and standardized mean difference (SMD), but mean difference, risk ratio, rate ratio, hazard ratio, proportion, etc. are also possible summary measures. In case of variability in reported effects, several formulas for converting test statistics (such as *t*-, χ^2-, *Z*- or *F*-values, or their associated *p*-levels) to effect size estimates (such as Cohen's *d*, OR, and correlation coefficients) and formulas for converting effect size estimators from one type to another are available (Lipsey & Wilson 2001, DeCoster 2004).

In the area of psychosomatic research, however, there may be specific problems concerning retrieving effect sizes, as different measures for the same construct are often applied in the original studies. For example, in a meta-analysis on depression and pain perception thresholds, effect sizes had to be calculated from studies using different methods to measure depression (Beck Depression Inventory, Hamilton Depression Scale, or diagnostic criteria according to Diagnostic and Statistical Manual, Fourth Edition) and pain perception (cold-, heat-, pressor-, ischemic-, and electrical stimuli) (Dickens et al. 2003). Effect sizes based on the standard deviation, such as the SMD, provide a solution to this problem. The SMD is calculated as the difference between the means of the cases and controls divided by the pooled standard deviation. Although the SMD is the predominantly used effect size in social sciences (Rosenthal & DiMatteo 2001), it has been disputed by others (Greenland & O'Rourke 2008). The use of SMDs is criticized because of the underlying assumption that the different scales from which the SMD is calculated are linearly related. Although this is a limitation of

the SMD, calculating SMDs is sometimes the only feasible option (e.g., in cross-sectional case-control studies).

Correlation coefficients and standardized mean differences (SMD)
When an effect size is applied to continuous variables, commonly used effect size indices are the *r* family and the *d* family (Rosenthal & DiMatteo 2001). The *r* effect size family includes all types of correlation coefficients (i.e., *r*, phi, rho, etc) and is preferably used when the studies composing the meta-analysis primarily report the correlation between variables (e.g., continuously measured psychological factors, such as anger or hostility, and development of hypertension in population cohorts (Rutledge & Hogan 2002)). The *d* effect size family provides SMDs and is preferably used when the studies composing the meta-analysis primarily report ANOVAs and *t*-tests comparisons between groups on continuous variables (e.g., cardiovascular activity in post-traumatic stress disorder patients versus healthy controls (Buckley & Kaloupek 2001)). Cohen's *d* and Hedges' *g* are two widely used examples of a SMD. When sample sizes are small, Cohen's *d* may produce estimates of the population effect size that are slightly too large (Hedges & Olkin 1985). Therefore, Cohen's *d* is sometimes adjusted by the following formula ($\sqrt{}$ [(*N* cases + *N* controls – 2) / (*N* cases + *N* controls)] * *d*), which results in Hedges' *g* (Hedges 1981). For example, a study compared cardiac vagal activity in 11 patients with chronic fatigue syndrome (2.75 SD 1.39 ln ms^2) with 11 healthy controls (3.09 SD 0.56 ln ms^2) (Cordero et al. 1996). It can be calculated that Cohen's *d* = -0.34 (95% CI = 0.00 to -1.14). In this small sample, correction leads to a small attenuation of the effect size, namely Hedges *g* = $\sqrt{}$ [(11+11-2) / (11+11)] * -0.34 = -0.32. The *r* of this association is 0.17 based on Cohen's *d* (r = $\sqrt{}$ [d^2 / [d^2 + 4], indicating that explained variance (r^2) chronic fatigue syndrome by lower cardiac vagal activity in this study is 0.17^2 = 2.8% (and 2.5% when using Hedges *g*).

Occasionally, descriptive or inferential statistics needed to compute an effect size are not reported at all. Conservative approaches to impute an effect size for missing values exist. For example, when a significant association was reported in the primary study, a conservative effect size assuming that the *p*-value was equivalent to 0.05 can be computed. In case there was no significant association, an effect size of 0.00 can be imputed. Several methods to impute missing data in meta-analyses have been discussed elsewhere (Wiebe et al. 2006).

The magnitude of effect sizes is often interpreted using Cohen's conventions, in which a SMD of 0.0 means no difference, 0.2 represents a small difference, 0.5 a moderate difference, and 0.8 a large difference (Cohen 1988). Inherent to the multifactorially caused conditions typically under investigation in psychosomatic research, effect sizes are usually small. For example, the median Cohen's *d* in the studies listed in Table 1 is 0.34 (interquartile range 0.27 - 0.55). The final evaluation of the meaning of the effect size, nevertheless, requires individual judgment regarding the specific topic under study, in which the consequence of

the outcome or the possibility of prevention and treatment are also taken into account.

Odds ratios (ORs) and other probability effect sizes
Probability effect sizes are usually given in studies with a binary outcome, such as in studies with disease versus no disease or mortality versus no mortality as endpoint. The selection of the appropriate summary statistic is a subject of debate due to conflicts in the relative importance of mathematical properties and the ability to intuitively interpret results (Sutton & Higgins 2008). Recommendations how to choose between ORs, risk differences, risk ratios, and other relative measures have been given elsewhere (Deeks 2002). Frequently, ORs are reported, such as in a meta-analyses on the association between depression and cardiovascular disease and mortality (Wulsin & Singal 2003, van Melle et al. 2004, Barth et al. 2004). The odds are the number of patients who fulfill the criteria for a given endpoint divided by the number of patients who do not. The OR relatively easily allows combining data and testing for statistical significance.

Although relative effect measures are generally used for summarizing the evidence, absolute measures, such as the absolute risk reduction or the number of patients needed to treat to prevent one event are more useful when applying the results in a concrete clinical or public health situation, and should be recalculated from the relative summary estimates (Egger et al. 1997b, Altman 2000). Also, effect sizes should be compared to the effects of well-established risk factors in the field in order to determine their (clinical) importance. For example, the influence of depression as a risk factor for the development of coronary disease in community samples without clinically apparent heart disease was considered similar to published effect sizes of the widely accepted risk factor smoking (Wulsin & Singal 2003).

Pooling of effect sizes
Methods for calculating the summary estimate by combining the effect sizes of the individual studies use a weighted average of the results, in which larger studies have more influence than smaller ones. This study weight is computed from the variance or squared standard error of the mean (SEM): weight factor (w) = 1 / SEM^2. The summary estimate in a meta-analysis is the mean weighted effect size, calculated by the sum of the products of effect size and weight per study, divided by the sum of all study weights ($\sum$ effect size * w / $\sum w$). The accompanying 95% confidence interval can be calculated with the following formula: $\pm$ 1.96 $\sqrt{}$ 1 / (SEM of summary estimate)^2. It is assumed that each value contributing to the summary estimate is statistically independent of the others. An extensive overview of the statistical basis of formal meta-analysis has been provided by others (Fleiss 1993). The meta-analysis can be repeated using

different methods to assess whether the same results are achieved and the summary estimate is robust to the decisions made to obtain it.

Software
Programs designed for the statistical pooling of data in meta-analysis are available, and most general statistical packages include meta-analysis options. Most of these programs are relatively easy to master and offer tutorials and a help function. Moreover, many programs and add-ons to statistical packages are freely available on the Internet. In general, programs offer at least basic statistical methods and graphical presentations and commercial software is not necessarily better than free software (Bax et al. 2007). Differences may exist in statistical methods, usability, graphics, and whether or not the software is being maintained. The basic results obtained from the different software packages are essentially the same (Egger et al. 1998b). For an overview and comparison of some meta-analysis programs, see for example Bax et al. (Bax et al. 2007) or Egger et al. (Egger et al. 1998b).

Fixed effect versus random effects models
When pooling the effects of all studies included in the meta-analysis, the fixed effect model or the random effects model can be used. The fixed effect model assumes that the samples of all studies are based on the same population and that the same underlying effect is thus measured in all studies (i.e., there is one true effect size). In this method, between-study variation is assumed to be due to sampling error. A disadvantage of the fixed effect model is that it is highly unlikely that studies do measure the same underlying effect, especially in epidemiologic research (Colditz et al. 1995). The random effects model, in contrast, assumes that each sample comes from a somewhat different population and that the effects in these populations may also differ. Between-study variation is assumed to be due to differences in the underlying effects in the samples. The random effects model gives the average effect of all studies (Higgins & Green 2008). A disadvantage of the random effects model is that it assumes the studies are a random sample of effect sizes and that between-study variation is distributed normally (Hardy & Thompson 1998). This is often not the case, for example as a result of publication bias (Colditz et al. 1995). An advantage of the random effects model is that it permits to generalize to studies that might be done in the future.

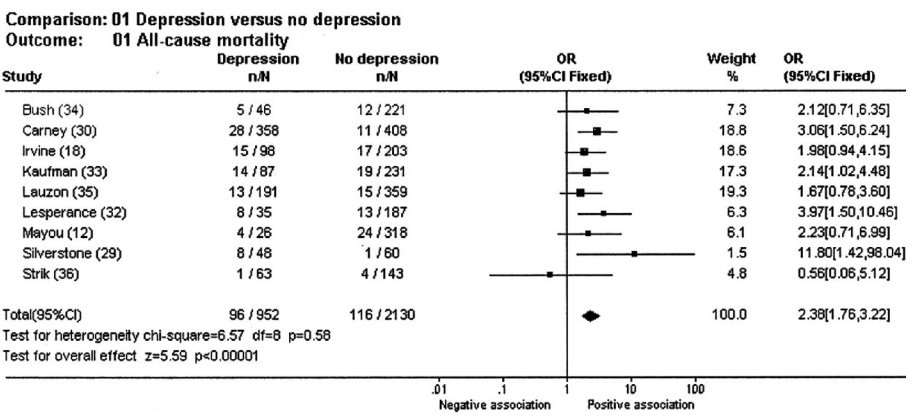
In the fixed effect model, the inverse variance method is used to pool effect sizes based on continuous data, such as mean differences or SMDs. To pool effect sizes based on binary data, such as ORs and RR's, the Mantel-Haenszel's method can be used (Mantel & Haenszel 1959), or the Peto method in case of pooling ORs of studies with balanced arm sizes, small intervention effects, or rare events (Peto et al. 1977). When the effect sizes are pooled using a random effects model, DerSimonian and Laird's method is used both for effect sizes based binary and continuous data (DerSimonian & Laird 1986).

The random effects model is more conservative than the fixed effect model and is used when heterogeneity is suspected (Egger et al. 1997a). Although tests for heterogeneity are often used to determine whether a fixed or random effects model must be used, these tests are often underpowered and deciding on the model should therefore be primarily based on characteristics of included studies (Hardy & Thompson 1998, Higgins & Green 2008). In general, the random effects model is more plausible and using the fixed effect model should only be done when this can be firmly justified on theoretical grounds. An example of how using random effects versus fixed effect analysis can change the summary estimate and conclusion is found in a meta-analysis on cortisol levels in patients with functional somatic syndromes (Tak et al. 2009d). The fixed effect model shows significantly lower cortisol in patients with functional somatic syndromes compared to healthy controls (SMD -0.12, 95% CI -0.18 to -0.05, $p<0.01$), whereas the more appropriate random effects model shows a wider confidence interval and no statistical significant difference (SMD -0.07, 95% CI -0.17 to 0.04, $p = 0.24$).

Forest plot

The main results of a meta-analysis are usually represented in a forest plot. Forest plots graphically display information on the individual studies included in the meta-analysis, the amount of variation between studies, and an overall estimate of the results of all studies combined (see Figure 1) (Borenstein et al. 2009). ORs are best plotted on logarithmic scales, as this enables ORs of the same magnitude but opposite directions - for example, 0.1 and 10.0 - to be equidistant from 1.0 (Egger et al. 1997b).

Figure 1. *Example of a forest plot on the prognostic value of depression on cardiovascular mortality. ORs are plotted on the logarithmic scale (van Melle et al. 2004). Reprinted with permission.*



Next to the forest plot, the basic details of each study supplying data should be presented, such as primary author, year of study, design, crude data, derived summary estimate and measure of its precision, allowing readers to evaluate the summaries against what was presented in original reports, or to repeat the meta-analysis while making other decisions or using other techniques.

Heterogeneity

Heterogeneity in meta-analysis means that included studies differ considerably on one or several important aspects, which may affect comparability of their results and which may have caused differences in results. Studies can be different in (a) biological, psychological, or clinical variables including gender, age, characteristics of study participants, severity of exposure, and condition or disease, (b) methodological variables including study design, measurement procedures, extent of control for confounding, and response measures, and (c) miscellaneous variables including year of publication, characteristics of the authors, and funding.

The presence of heterogeneity can be calculated statistically (Higgins & Thompson 2002b). The most used measures are the Q statistic, I^2, and tau-squared (τ^2). To begin with, the Q statistic (also called Cochran's χ^2 statistic) is a chi-squared test calculating whether variation in study results is due to chance or that variation is due to systematic underlying differences and the null hypothesis should be rejected. A value of Q similar to the degrees of freedom in the analysis indicates little heterogeneity. When it is considerably higher, and the p-value is lower than 0.10, this indicates heterogeneity (Thompson 1994). The Q statistic, however, has a number of limitations. The Q statistic has low power when a single study largely contributes to the mean weighted effect size (Hardy & Thompson 1998). It also has low power when included studies are small or when there are few studies, whereas this test may detect heterogeneity even when it is not substantial when many studies are included (Higgins & Green 2008). In the studies presented in Table 1, meta-analyses in which the Q test was not significant were usually based on a small number of primary studies (less than 10), whereas the Q test was significant in all meta-analyses with a relatively large number of included primary studies (more than 30). This implies that heterogeneity can be considered the rule, rather than the exception, in meta-analyses published in Psychosomatic Medicine. A second measure of heterogeneity is the I^2 statistic, which is a derivative of Q. This statistic gives the percentage of variability in results that is caused by heterogeneity rather than coincidence (Higgins & Green 2008). Generally, an I^2 of over 50% indicates considerable heterogeneity. A third statistic that is often used to report heterogeneity is τ^2. This is the variance of the true effect size, thus, there is no heterogeneity when this statistic is 0 (Colditz et al. 1995). There are several more statistics that assess heterogeneity, some specific to the type of effect measure (Hardy & Thompson 1998).

However, because of problems with power and accuracy, when statistical tests of heterogeneity indicate that the null hypothesis of no heterogeneity holds, this does not indisputably prove that studies are completely homogeneous (Hardy & Thompson 1998). It should still always be investigated whether studies have important clinical and methodological differences. When heterogeneity is suspected, this must be accounted for in the statistical analysis and mentioned in the discussion. Usually the presence of heterogeneity is considered to be a negative aspect of meta-analysis, because it makes the results of the meta-analysis difficult to interpret (Hardy & Thompson 1998) and suggests that samples may be too different to be combined. However, as no widely accepted quantitative measure exists to grade heterogeneity, it may be better to examine it in a meta-analysis rather than use it as a reason for not conducting one (Ioannidis et al. 2008). Heterogeneity can also have advantages. If studies that are clinically and methodologically heterogeneous lead to comparable results, this means that the results are generalizable to a wider population. In addition, investigating sources of heterogeneity can lead to a better understanding of associations, new hypotheses (Counsell 1997, Hardy & Thompson 1998), and improvement of future research (Berlin 1995).

Moderator analysis

Performing meta-analysis on subgroups based on characteristics that potentially are responsible for differences in effect sizes between studies, can demonstrate whether the strength of the summary estimate is influenced by these characteristics. This procedure is referred to as moderator analysis. For example, in a meta-analysis a significant difference was found in interleukin-6 (IL-6) serum levels between depressed patients and controls, with a SMD of 0.25, 95% CI 0.18-0.31. However, the magnitude of the summary estimate of the association between IL-6 and depression was largely attenuated in studies that adjusted for BMI (N = 22, SMD = 0.08, 95% CI 0.02 to 0.13, p = 0.007) as compared to studies without BMI adjustments (N = 39, SMD = 0.50, 95% CI 0.37 to 0.63, p<0.001) (Howren et al. 2009).

Ideally, such subgroup analyses are planned in advance, since investigating heterogeneity post hoc based on the data from the meta-analysis itself can lead to over inclusion (Hardy & Thompson 1998). Providing a rationale for each moderator and giving due consideration to the role that each moderator is intended to play is essential (Freedland et al. 2009). In case of post hoc subgroup analyses, results should be reported as exploratory and the need for replication should be mentioned (Colditz et al. 1995). Particularly in observational studies, possible moderator analyses on confounders, moderators, and mediators are an important part of the meta-analysis. The extent to which putative confounders, moderators, and mediators have been taken into account in original studies is often highly variable and extracting useful data is not always possible.

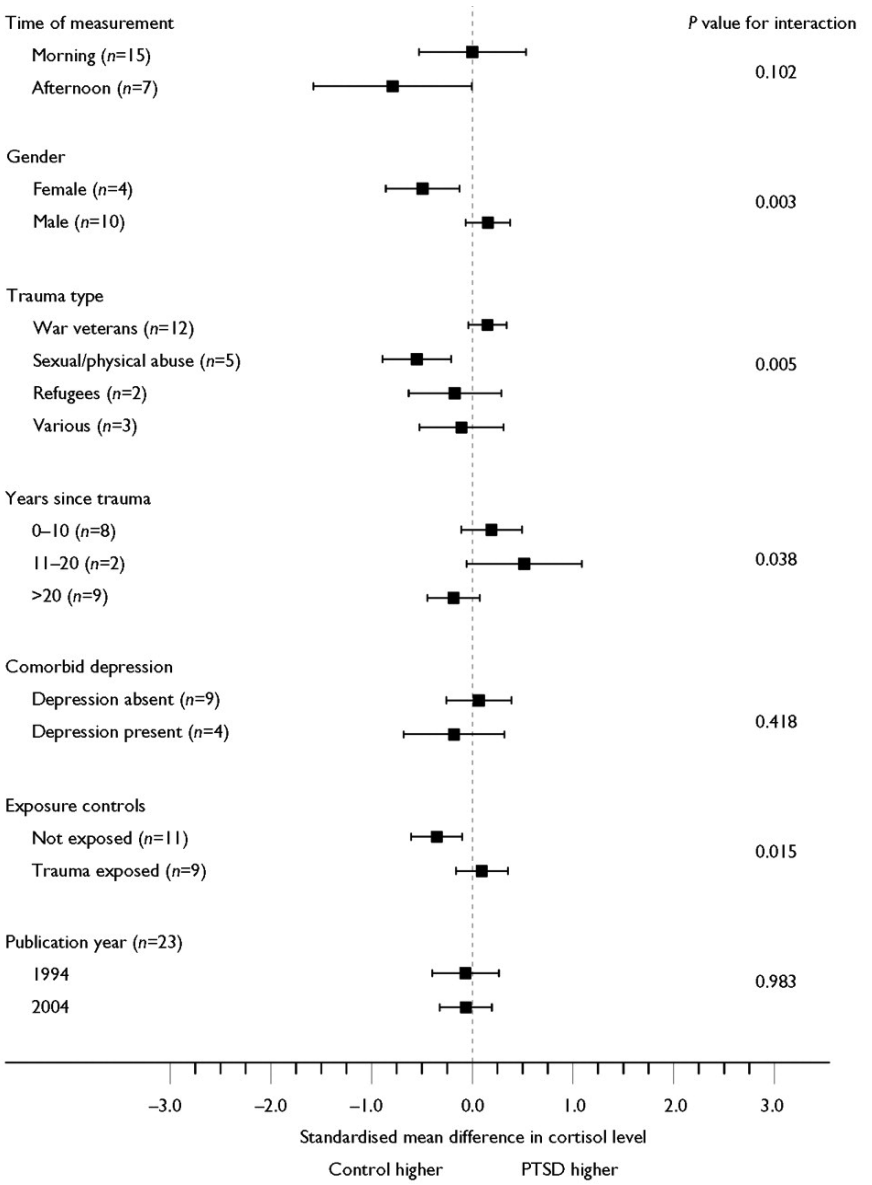
Some difficulties may arise when using terminology regarding confounders, moderators, and mediators. A variable may be considered as a confounding or mediating factor in the original study, but this variable is tested as a moderator in the meta-analysis. For example, authors in the previously mentioned meta-analysis on IL-6 and depression propose that depressive symptoms may facilitate weight gain over time as a result of physical inactivity. In this pathway, BMI may be a mediator in reality (in case depression leads to weight gain and weight gain to inflammation), but is referred to as a moderator of the effect size in the meta-analysis.

The difficulties faced in moderator analyses are many. First, there is the risk of spurious findings due to multiple testing. When the number of original studies in the meta-analysis is small (i.e., N = 10 - 15), there are insufficient degrees of freedom to test more than one moderator variable (Babyak 2004). Nevertheless, many more subgroup analyses are often performed, as illustrated by some of the studies listed in Table 1. Second, when moderating variables are continuous they have to be categorized in order to be able to perform a moderator analysis. It is, however, often unknown how to define subgroups. Artificially grouping data into categories introduces measurement error with an inevitable loss of power (Freedland et al. 2009). Furthermore, the arbitrariness of the choice of cut point may lead to the undesirable temptation of trying more than one value and choosing the one that gives the most satisfactory result (Altman 2000). Third, moderator analysis does not provide a statistical test of the existence of a moderator effect. Fourth, one cannot look at effect moderation while keeping other covariates constant. When two moderators are highly correlated and the first causes changes in the effect size, a moderator test for the second will likely also be significant, even though this second moderator does not truly influence the strength of the effect.

An example of difficulties in interpreting the results of moderator analysis is a meta-analysis across 37 studies on cortisol levels in patients with post-traumatic stress disorder (PTSD) (Meewisse et al. 2007). Overall, cortisol levels were not significantly different in PTSD patients compared to controls (SMD -0.12, 95% CI -0.32 to 0.08, p = 0.24) (Meewisse et al. 2007). Figure 2 presents subgroup analyses (SMD and 95% CI are shown) based on several moderators. Although there is no significant difference in cortisol between PTSD patients and controls in the primary summary estimates, significantly lower levels of cortisol are found in females with PTSD compared to female controls (gender is a moderator), in patients with PTSD with physical or sexual abuse compared to controls (trauma type is a moderator), and in PTSD patients when they are compared to controls without trauma exposure, as opposed to controls with trauma exposure but without PTSD (exposure status of control group is a moderator). However, as the authors mention, it is not possible to disentangle whether those moderators act separately from each other. For example, the association between female gender

and lower cortisol could be explained by a larger prevalence of sexual abuse in women. In this case, meta-regression could be a possible solution to further elucidate the independent contribution of those factors.

Figure 2. *Examples of moderator analyses in a meta-analysis on cortisol and post-traumatic stress disorder (Meewisse et al. 2007). Reprinted with permission.*



Meta-regression

Meta-regression is a regression-based analysis that aims to test for study heterogeneity by relating study characteristics to study outcome. Typically, the independent variables (predictors) are characteristics of each study, such as participants' mean age, proportion of women, or follow-up duration. The dependent (outcome) variable is the study effect size such as the SMD or log OR. The procedure for multivariable meta-regression closely follows conventional regression analysis, the only difference being that a variable equal to the inverse variance (i.e., the study weight) has to be used as case weight in order to perform a weighted regression. Meta-regression can be used to explain heterogeneity and provides the possibility to simultaneously assess multiple characteristics. Again, the fixed effect or random effects model can be used for meta-regression. The full range of regression models and methods (i.e., linear or logistic regression, testing interactions, model fitting statistics, etc) can be employed (Greenland & O'Rourke 2008). For example, in a meta-analysis on placebo response in chronic fatigue patients, it was found that the placebo response was higher in interventions based on immunological assumptions compared to interventions based on psychological assumptions. The authors hypothesized that this difference could be explained by higher expectations of patients on interventions assuming physical causation as opposed to interventions assuming psychological causation. Alternatively, they also considered the possibility that systematic differences between immunological and psychological trials, such as illness duration, placebo type, and duration of follow-up, could explain the larger placebo effect in immunological trials. In a meta-regression, however, only intervention type (i.e., psychological or immunological) set out to be significantly associated with a stronger placebo response ($p = 0.03$), independent from all other factors (Cho et al. 2005).

Some problems affect the validity and reliability of meta-regression. Primarily, meta-regression is prone to inflate false-positive rates when heterogeneity is present, when there are few studies, and when there are many covariates. Consider the case of two studies producing effect estimates with non-overlapping confidence intervals: any covariate whose value differs between these studies will be significantly related to the heterogeneity among the studies, and hence a potential explanation of it, while this explanation could be entirely spurious (Higgins & Thompson 2004). Furthermore, it is unclear how many covariates can reliably be investigated without the risk of overfitting, and how this depends on the number of studies, the extent of the heterogeneity, and the relative weights of the different studies. Rules of thumb for conventional regression analyses (10 - 15 observations per covariate assessed, for instance (Babyak 2004)) are not directly relevant to meta-regression, as this type of regression deals with the complexities of heterogeneity and differential study weights. To be on the safe side, meta-analysts who aim to explore heterogeneity using meta-regression should

minimize the number of covariates investigated, select those justified through scientific rationale, and specify them in advance (Higgins & Thompson 2004). Second, regarding translation of the results of the meta-regression to individuals, the problem of aggregation bias ('ecological fallacy') may arise. This bias refers to the assumption that individuals have the average characteristics of the group to which they belong, and thus that relationships observed for groups necessarily hold for individuals. The meta-regression analysis is conducted at the study-level and does not include the underlying patient-level variation. The relationship between group means may not reflect the relationship between values of exposure and outcome in an individual, and average quantities instead of person-specific quantities can lead to erroneous conclusions (Schwartz 1994). For example, suppose that countries with a high per capita income have high suicide rates. Inferring that increasing personal income at the individual level is also associated with suicide related mortality can lead to erroneous conclusions, as within countries, suicide related mortality may be lower in high income than in low income persons. Thus, when interesting findings are discovered using meta-regression, person-level data from large cohort studies or trials may be required for confirmation.

Alternatively, meta-analysis of individual patient data (IPD) (also referred to as 'mega-analysis') could be considered, in which raw data from every primary study are obtained and transformed to a common format. The strengths of IPD meta-analysis in general are that the power is greatly enhanced by the larger number of subjects and that more subgroup analyses can be done. More importantly, however, in observational research, individual patient data can be used to consistently adjust for confounders. Adjustment for confounders is usually impossible in common meta-analysis, as not all studies perform the same adjustments in their analyses and they report adjusted analyses in different ways. An example is the meta-analysis on the impact of depression on mortality. In this research, several moderator analyses, such as on measurement instrument to assess depression or duration of follow-up, did not explain heterogeneity (Barth et al. 2004). Also, the relative risk of mortality was nearly the same in unadjusted and adjusted results, and the amount of heterogeneity was not reduced. Authors argue that one possible explanation for the heterogeneity of the adjusted analyses may be the selection of risk factors, which varied greatly from study to study. One possible solution to this problem would be to pool and reanalyze the original data of all included studies. This can only be done when different studies include comparable measures of the variables to be adjusted for. A major obstacle of IPD meta-analysis is that it is time-consuming and requires cooperation between several research groups, which may not always be attainable. In addition, variables that must be compared will generally be measured using different instruments in the individual studies and must therefore be harmonized before analysis is possible. Information may be lost during this process of

harmonization. A good overview of the methodology of IPD meta-analysis is given by Stewart et al. (Stewart & Clarke 1995).

INTERPRETATION OF META-ANALYSIS FINDINGS

Sensitivity analysis

The process of undertaking a meta-analysis inevitably involves many more or less subjective decisions and sensitivity analyses can be carried out to determine whether the assumptions or decisions made have a major effect on the result of the meta-analysis. Thus, sensitivity analysis addresses the question whether the findings of the meta-analysis are robust to the methods used to obtain them. Examples of sensitivity analyses include assessing the influence of including studies that were doubted to meet eligibility criteria, comparing fixed effect with random effects models, comparing cohort and case-control studies, adding conservative effect size estimations for studies that did not provide adequate data to calculate an effect size, or excluding outlying studies. Two other commonly performed sensitivity analyses are assessing the influence of methodological study quality of primary studies and the influence of publication bias.

Using methodological quality in interpreting meta-analysis results

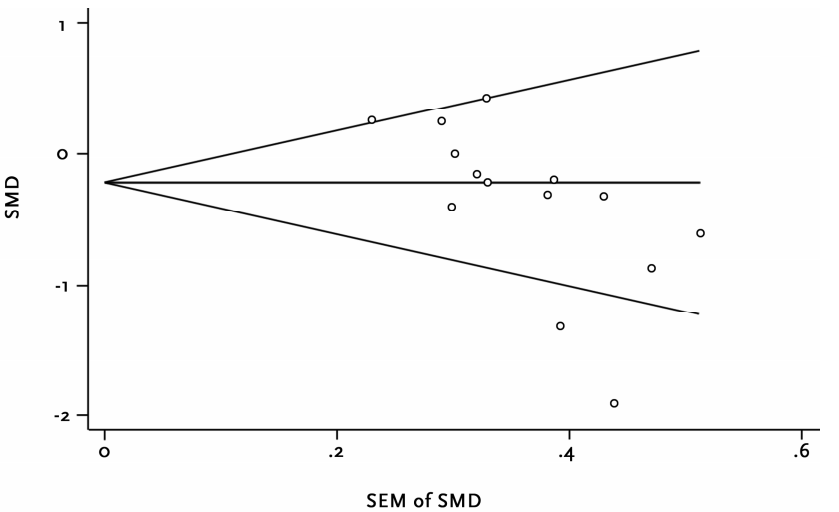
It remains a matter of debate how the results of quality assessment should be incorporated in the analysis and interpretation of results of meta-analyses. Exploring the effects of quality on the quantitative results by using quality as a weighting factor has been discouraged (Detsky et al. 1992). We recommend using quality scores in a sensitivity analysis, which can demonstrate whether the findings of the meta-analysis are different for low and high quality studies. For example, sensitivity analysis in a meta-analysis on the association between positive well being and mortality in healthy populations indicated a stronger association between positive psychological well-being and reduced mortality in high quality studies compared to low quality studies (Chida & Steptoe 2008b). This sensitivity analysis thus supports the validity of the overall finding that there is an association.

Publication bias

Publication bias in observational meta-analyses may lead to inflated effect estimates that tend to be in the hypothesized direction. Several approaches have been developed to assess publication bias. The most well-known approach is the funnel plot: a scatter graph with for each primary study the effect estimate plotted against a measure of precision (such as sample size, or preferably, the standard error of the effect size) (Sterne & Egger 2001). It is expected that more precise studies report effect estimates close to the true effect, whereas effects estimates from less precise studies will scatter more widely. In the absence of publication

bias, the plot is expected to resemble a funnel-like shape, which is symmetrical around the summary estimate. In a meta-analysis on decreased cardiac vagal activity in functional somatic syndromes, the funnel plot was not symmetric, as there is gap where small studies with null findings are expected (Figure 3A).

Figure 3A. *Funnel plot (N = 14) showing the correlation between the standardized mean difference (SMD) and its standard error (SEM) with pseudo 95% confidence limits. The summary estimate reveals that cardiac vagal activity is significantly lower in patients with functional somatic disorders compared to healthy controls (SMD = -0.32, 95% CI -0.63 to -0.01, p = 0.04).*



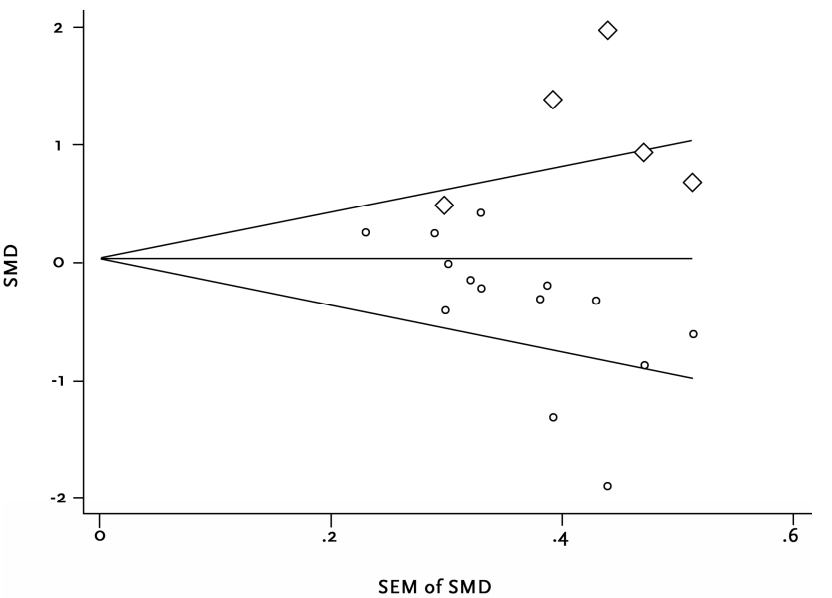
Funnel plots can be visually interpreted, but this is subjective and the agreement between raters and the association between graph ratings and publication bias is found to be poor (Bax et al. 2009). A test for funnel plot asymmetry formally examines whether the association between estimated effects and a measure of study precision is larger than might be expected to occur by chance. The principle is to relate the effect estimates to their SEM, and to test the null hypothesis that the association is absent. There are many tests for funnel plot asymmetry, which are compared by Rucker et al. (Rucker et al. 2008). Two of the most well-known are the Begg and Mazumdar adjusted rank correlation test, and the Egger et al.'s regression asymmetry test. The Begg and Mazumdar test (Begg & Mazumdar 1994) is based on a Kendall's τ rank correlation between the standardized effect size and its SEM. Egger's test (Egger et al. 1997a) is based on a linear regression of the effect estimate against its standard error, weighted by the inverse of the variance of the effect estimates. However, the tests have low power to detect funnel plot asymmetry, and do thus not exclude the presence of publication bias.

Publication bias is not the only reason for funnel plot asymmetry. Asymmetry also arises because of small-study effects - a tendency for the effects estimated in smaller studies to differ from those in larger studies (Sterne et al. 2000). Small study effects occur because of differences in methodological quality between larger and smaller studies, heterogeneity between studies with different sample sizes (small studies may more likely include selected groups of patients), an effect modifier associated with study precision, or merely chance (Egger et al. 1997a, Sterne et al. 2000). In addition, some effect estimates (e.g., ORs and SMDs) are naturally correlated with their standard errors, and can produce spurious asymmetry in a funnel plot (Higgins & Green 2008). Another mathematical estimation of publication bias is provided by the fail safe *N*, which indicates the number of new, unpublished, or unretrieved non-significant studies that would be required to lower the significance of a meta-analysis to non-significant. A fail safe *N* that is small, particularly compared with the number of studies included in the meta-analysis, indicates that the degree of confidence that can be placed in the main conclusions of the meta-analysis is low. The fail safe *N* has been criticized for two reasons. Firstly, it overemphasizes statistical significance. Secondly, it is based on the addition of studies that have an average null effect, whereas unpublished studies may also have an effect in the opposite direction as the observed meta-analysis result (Evans 1996).

An important question is how to proceed when publication bias is suspected. A relatively simple approach to correct for publication bias is the 'trim and fill' method (Duval & Tweedie 2000a, Duval & Tweedie 2000b), available in most statistical meta-analysis programs. The principle behind this method is to impute new studies to an asymmetric funnel plot, followed by a meta-analysis that includes the imputed studies. The method works by estimating the number of studies on the right-hand side of the funnel plot that have no counterpart on the left-hand side. Studies causing the asymmetry are then 'trimmed' from the right-hand side of the funnel plot, possibly leading to a shift of the re-estimated summary estimate that may again create asymmetry. The process is repeated until there is no residual asymmetry, after which the trimmed studies are put back and their missing counterparts are imputed or 'filled' by replicating the opposite side of the funnel plot with the mirror axis placed along the adjusted summary estimate. The difference between the original summary estimate and the summary estimate based on the extended dataset including the imputed studies is assumed to indicate the degree of publication bias. Assumptions underlying the trim and fill method include that the magnitude of the effect size, and not the *p*-value, determines the chance of publication. Moreover, this technique assumes that publication bias leads to this simple form of funnel plot asymmetry, and that missing effect size estimates are of the same size as those observed in the opposite direction. Nevertheless, it has been shown that the use of the trim and fill method can help to reduce the influence of publication bias on the summary

estimates, even though the performance of this method decreases when heterogeneity increases (Peters et al. 2007). With regard to the studies on cardiac vagal activity in functional somatic syndromes, Egger's test rejected the null hypothesis that there was no funnel plot asymmetry ($p = 0.01$). The trim and fill method resulted in a fill of five studies and a shift from an initially significant summary estimate to a re-estimated summary estimate that was non-significant (Figure 3B) (Tak et al. 2009a). This analysis points to the possibility that studies contradicting prevailing beliefs of lower cardiac vagal activity in functional somatic syndromes have not been published. The trim and fill method is recommended to be used as a form of sensitivity analysis of the summary estimate.

Figure 3B. *Trimmed and filled funnel plot (N = 19) showing the correlation between the standardized mean difference (SMD) and its standard error (SEM) with pseudo 95% confidence limits. Squares represent the studies that have been filled. The adjusted summary estimate reveals that cardiac vagal activity is not significantly different in patients with functional somatic disorders compared to healthy controls (SMD = 0.01, 95% CI -0.34 to 0.36, p = 0.95).*



Controversy around meta-analysis

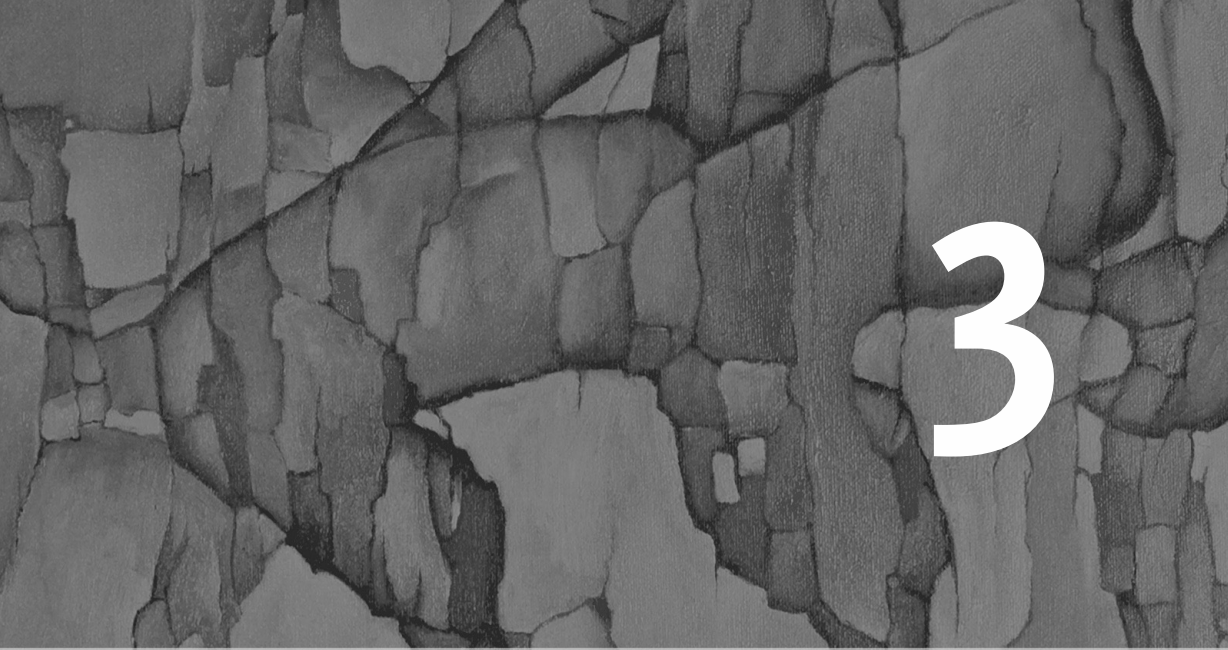
There are a number of outspoken critics of meta-analysis. Most points of criticism do not only apply to meta-analysis, but to the entire field of observational research, such as the risk of reporting bias, publication bias, confounding, and

lack of comparability between studies. Some even argue that meta-analysis of observational studies should not be done at all, because it would only reinforce the biases inherent to epidemiologic research by creating significant but incorrect results (Shapiro 1994). In a properly performed meta-analysis, however, these limitations can be dealt with in a sound way, as has been discussed in this article. Some critics argue that the statistical pooling of data in observational data is highly prone to bias and spurious findings. Instead, it is suggested that it is more important to thoroughly investigate causes of heterogeneity (Egger et al. 1998a). We agree that statistical combination of studies should not generally be the main aim of systematic reviews of observational studies, especially as heterogeneity seems the rule rather than the exception (Table 1). The thorough consideration of possible sources of heterogeneity between studies, by using moderator analysis, meta-regression, and sensitivity analysis should be considered as more important features of meta-analysis in this field. When there are still serious limitations to the results of the meta-analysis, these can be discussed and interpretation can be adjusted accordingly. Thus, instead of disputing the technique of meta-analysis itself, we feel its unduly aura of providing the final answer should be rectified.

CONCLUDING REMARKS

Many papers in Psychosomatic Medicine primarily are observational studies aiming to answer etiological questions. Apart from providing a summary estimate, the importance of meta-analyses based on those studies also lies in the identification of sources of bias, heterogeneity, generation of new hypotheses, and the construct of guidelines to conduct better research in the future. Rather than pretending to provide the final, not debatable answer, meta-analysis relies on shared subjectivity. Every analysis inevitably requires certain subjective decisions, but these have to be transparent and explicit. The discussion of a meta-analysis should not simply state the results of the statistical pooling, but it should also discuss the level of certainty of the conclusions and any limitations to the interpretation of the findings. Specific guidelines on adequate reporting of meta-analyses based on clinical trials (i.e., PRISMA) or on observational studies (i.e., MOOSE) are available (Stroup et al. 2000, Moher et al. 2009).

This review aimed to demonstrate that performing a meta-analysis is a good way to gain more knowledge concerning a specific research topic. We agree with Rosenthal et al. when they state that anyone who is considering a review of the literature has little justification for not doing it quantitatively (Rosenthal & DiMatteo 2001), as the skills and training required for performing a high quality meta-analysis are modest. We hope that the number of meta-analyses in Psychosomatic Medicine will increase, as they have the ability to produce more knowledge than is provided by the sum of its parts.



3

CHAPTER 3

As good as it gets?
A meta-analysis and systematic review
of methodological quality of heart rate variability studies
in functional somatic disorders

LM Tak, H Riese, GH de Bock, A Manoharan, IC Kok, and JGM Rosmalen
Biological Psychology 2009;82(2):101-11

ABSTRACT

Autonomic nervous system (ANS) dysfunction is a potential mechanism connecting psychosocial stress to functional somatic disorders (FSD), such as chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome. We present the first meta-analysis and systematic review of methodological study quality on the association between cardiac ANS dysfunction, measured as parasympathetic nervous system (PNS) activity using heart rate variability (HRV), and FSD. Literature search revealed 23 available studies including data on 533 FSD patients. Meta-analysis on a subgroup of 14 studies with suitable outcome measures indicated lower PNS activity in FSD patients compared to controls (weighted standardized mean difference (SMD) = -0.32, 95% CI -0.63 to -0.01, p = 0.04). The reliability of this summary estimate was, however, significantly limited by unexplained heterogeneity in the effect sizes and potential publication bias (weighted SMD after correction for funnel plot asymmetry = 0.01, 95% CI -0.34 to 0.36, p = 0.95). The systematic review of overall methodological quality of HRV studies in FSD demonstrates that there is substantial room for improvement, especially in selection of healthy control subjects, blinding of researchers performing HRV measurements, report of adequate HRV outcomes, and assessment of and adjustment for potential confounders. Methodological study quality was, however, not a significant predictor of study findings. We conclude that current available evidence is not adequate to firmly reject or accept a role of ANS dysfunction in FSD. Quality criteria and recommendations to improve future research on HRV in FSD are provided.

INTRODUCTION

Functional somatic disorders (FSD) are syndromes of related complaints with no known underlying organic pathology. The main three are chronic fatigue syndrome (CFS), fibromyalgia (FM) and irritable bowel syndrome (IBS) and it has been thought that their similarities outweigh their differences (Wessely et al. 1999). An important theory about the etiology of FSD is the chronic stress hypothesis: stress-sensitive personality traits, stressful life events, and chronic difficulties have part in the development of FSD (van Houdenhove et al. 2005, Deary et al. 2007).

Dysfunction of the autonomic nervous system (ANS) is a potential mechanism that connects psychosocial stress to FSD. It has been shown that the ANS is influenced by intense acute, repetitive, and chronic psychosocial stress (Sloan et al. 1994, Dishman et al. 2000, Schommer et al. 2003). When the load of stressors in an individual is too large, when the ANS is chronically addressed, or when the capacity of the ANS to adjust is diminished, ANS dysfunction may develop (Chrousos & Gold 1992). As the ANS innervates several organs, sensation and misinterpretation of generated peripheral physiological arousal due to dysfunction may contribute to experience of functional somatic symptoms (Sharpe & Bass 1992, Rief & Barsky 2005). A narrative review of 34 studies suggests that decreased parasympathetic nervous system (PNS) activity could be associated with presence of CFS, FM, and IBS, however, findings are not fully consistent (Tak & Rosmalen 2007). It remains to be elucidated why it seems that the severity of PNS alterations varies both within and among the three main FSD. Disparities in methodological study quality, such as biased selection of control subjects, no exclusion of cardio-active medication use, inaccurate PNS activity measures, and lack of adjustment for potential confounders may have contributed to the mixed findings.

The aims of the current study are to assess whether FSD are characterized by alterations in PNS activity and to examine whether this association is moderated by the type of FSD. Furthermore, we aim to study whether methodological study quality correlates with study findings. Usually, cardiac ANS function is non-invasively assessed by measurement of heart rate variability (HRV), a measure of interbeat interval fluctuations in heart rate. Commonly used indices are frequency domain and time domain measures. Due to unstandardized environmental influences, ambulatory HRV recordings during 24-hours are not comparable with those of short duration in a situation (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996). Reliability of HRV measurements performed during interventions such as tilt or pharmacological stimulation is generally poorer than when measured at rest (Sandercock et al. 2005b). Therefore, we selected only studies reporting on resting PNS activity in the frequency domain (HRV-HF), defined at 0.15 - 0.40 Hz,

or in the time domain (the root mean square of differences between successive interbeat intervals (RMSSD)), both recommended as preferential PNS measures (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996). Short term HRV-HF and RMSSD have a clear physiological basis and are highly correlated (Berntson et al. 1997). We present the first meta-analysis and systematic review of methodological study quality on the association between resting HRV and FSD.

METHODS

Search strategy

Relevant articles were identified by searching the databases of Medline (1966 - June 2007), Embase (1980 - June 2007), PsycINFO (1960 - June 2007) and Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1982 - June 2007). For searching Medline with the PubMed interface, a search string was formulated. The first component consisted of the FSD chronic fatigue syndrome, fibromyalgia, irritable bowel syndrome, and synonyms. The second component consisted of the terms autonomic nervous system, parasympathetic nervous system, heart rate variability, and synonyms. Searching Embase, PsycINFO, and CINAHL, terms included in our search were adapted according to the thesaurus of the respective database and explosion of the search terms was applied. References of original articles and related reviews were hand searched for additional citations. In a further effort to locate published and unpublished studies, authors who had published three or more papers on the subject of HRV in FSD were consulted for potentially missed results.

Screening and selection procedure

First, title and abstract of the articles were screened on inclusion criteria (case-control studies; cases are adults with CFS, FM or IBS; measurement of ANS using HRV) by two independent reviewers (L.M.T. and J.G.M.R.). Discrepancies were resolved by consensus. Subsequently, full text articles were acquired and screened on the following exclusion criteria: a) no translation available, b) double report on same subjects, c) also cases with sub threshold FSD included, d) no outcome of HRV-HF or RMSSD reported, e) only 24-hour measurement of HRV, and f) HRV not measured in a resting condition.

Data extraction

Name of first author, year of publication, type of FSD, number of participants, age of the participants, and outcome regarding resting HRV were extracted from every paper. When several HRV measurements in different resting postures (supine or sitting) were reported, results in the supine position were preferred. If data were not extractable, i.e., no appropriate units and measures of dispersion or test

statistics were provided, authors were contacted and asked for additional information (Martinez-Lavin et al. 1997, Yataco et al. 1997, De Becker et al. 1998, Adeyemi et al. 1999, Orr et al. 2000, Thompson et al. 2002, Yamamoto et al. 2003, Waring et al. 2004, Robert et al. 2004, Cordova 2005, van Orshoven et al. 2006, Tousignant-Laflamme et al. 2006, Ng et al. 2007). Four authors supplied additional outcome information that was not published in the original publication (Yamamoto et al. 2003, Cordova 2005, van Orshoven et al. 2006, Tousignant-Laflamme et al. 2006). When efforts in making contact were unsuccessful or authors reported that data were not available, results were estimated from figures (Orr et al. 2000, Elsenbruch & Orr 2001b, Thompson et al. 2002, Ng et al. 2007). We decided to allow for different outcome measures (absolute values, logarithmically transformed values, and normalized units of HRV-HF) on PNS activity in FSD patients compared to controls. Additionally, we tabulated author conclusions, which are usually based on *p*-values. Meta-analysis was only performed on studies reporting absolute values or logarithmically transformed values and not on studies only reporting normalized units in accordance with an earlier meta-analysis (Rottenberg 2007). For each study, we calculated a standardized mean difference (SMD) of PNS activity as measured with HRV, which is an effect size measure based on the difference between mean values of patient and control groups divided by the pooled standard deviation (also referred to as Cohen's *d*) (Rosenthal & DiMatteo 2001). SMDs can be directly calculated for studies reporting logarithmically transformed values, which are usually normally distributed. To calculate a SMD of the highly skewed absolute HRV values, however, we first applied a formula to transform nonparametric data on the raw scale (Higgins et al. 2008). In case it was stated that a standard deviation was reported whereas it was more likely a standard error given its magnitude (Cohen et al. 2001b, Tousignant-Laflamme et al. 2006), the one which seemed most sensible was chosen (Deeks 1997). All data extraction was done by two independent reviewers (L.M.T. and J.G.M.R.).

Quality assessment

We defined a high quality study as a study with a design that minimizes bias in the estimation of the association between ANS dysfunction and FSD. To assess methodological study quality, we have searched for a reliable, validated quality tool without satisfactory result. Recently, this lack of a single tool for assessing quality of observational epidemiological studies has been issued (Sanderson et al. 2007). Therefore, we selected nine essential items from guidelines or tools for either reporting (von Elm et al. 2007) or appraising (Altman & Lyman 1998, Siegfried et al. 2003) observational research. We modified them for cross-sectional case-control studies of HRV in FSD and classified them according to three key domains (appropriate selection of participants, appropriate quantification of HRV, and appropriate control for confounding). In order to create a generally agreed standard, we consulted a panel of experts in the field

having three or more publications on the topic of HRV in FSD. Experts were allowed to consult or refer us to a colleague. Experts were asked to criticize each item, to advise about additional items which were missing, and to make a statement whether they approved the overall quality tool as an instrument for quality measurement in HRV studies in FSD. The response rate was 100% and experts of all three FSD (dr. S. Elsenbruch, prof. W. Orr, prof. M.M. Heitkemper, dr. R.L. Burr, dr. K. Cain, dr. M. Martinez-Lavin, and dr. K. Yoshiuchi) approved our quality tool. Taking their comments into account, a final version of the quality tool was constructed (see also Table 1).

Key domain 1 (items 1-4): Appropriate selection of participants
Patients have to meet the international consensus criteria for the respective FSD available at the time of the study (Center for Disease Control and Prevention Criteria for CFS, American College of Rheumatology Criteria for FM, Rome Criteria for IBS). Because self-report may introduce bias and underlying pathology has to be excluded, a physician is required to diagnose a FSD. Controls have to represent the population from which the cases arose. Recruiting healthy controls from hospital staff or medical students may lead to selection bias (Lee et al. 2007). A number of somatic conditions, such as diabetes mellitus (Malpas & Maling 1990), hypertension (Pagani et al. 1984), cardiovascular disease (Thayer & Lane 2007) and psychiatric conditions, such as depressive disorder (Rottenberg 2007) and anxiety disorders (Blechert et al. 2007, Friedman 2007), are known to compromise PNS activity. Additionally, several more rare diseases are associated with ANS alterations, such as chronic renal failure and some neurological disorders (van Ravenswaaij-Arts et al. 1993). Medication use is important to consider in HRV studies, as many medications act directly or indirectly on the ANS, such as antihypertensives (Schroeder et al. 2003), antidepressants (Rechlin 1994), antihistaminics (Nault et al. 2002), antipsychotics (Cohen et al. 2001a), and benzodiazepines (Agelink et al. 2002). Length of disease is difficult to assess reliably (e.g., formal diagnosis is typically much later than onset of symptoms), however, in the initial phase of the FSD other features may be present than in the chronic course of the disorder. Studies should report the central tendency of disease duration with an appropriate measure of dispersion. Since there is currently no tool that has been validated to indicate severity of the FSD, prevalence of diagnostic criteria should be reported as best available measure of FSD severity.

Key domain 2 (items 5-7): Appropriate quantification of HRV
Blinding of researchers helps to ensure reliability of study conclusions, which is not only important in clinical trials, but also in experimental research (Day & Altman 2000). Although unintended, researchers may approach FSD patients differently than healthy controls during HRV measurements. There is a wide range of other requirements for a valid HRV measurement and analysis; for technical details and an overview of requirements see the Task Force paper (Task Force of

Table 1. *Quality tool to assess methodological quality of heart rate variability studies in functional somatic disorders.*

APPROPRIATE SELECTION OF PARTICIPANTS

1- Has the disease of the cases been reliably assessed and validated?
- According to international criteria by a physician (2)
- According to international criteria, assessor not clearly established (1)
- Self report or not clearly stated (0)

2- Have all controls been recruited from the same population as the cases?
- Controls recruited from same population as cases (2)
- Selected population, such as hospital staff or students (1)
- Not clearly stated (0)

3- Is the population defined with in- and exclusion criteria?
- Medication use, somatic morbidity, psychiatric morbidity, 3 stated (2)
- Medication use, somatic morbidity, psychiatric morbidity, 1-2 stated (1)
- None stated or not clearly stated (0)

4- Are disease characteristics presented (length and severity of functional somatic disorder)?
- Duration of disease and severity of disorder is stated (2)
- Only duration or only severity is stated (1)
- None stated (0)

APPROPRIATE QUANTIFICATION OF HEART RATE VARIABILITY (HRV)

5- Is assessor of HRV blind for disease status?
- Yes (2)
- Not clearly stated (0)

6- Are methods for assessment of HRV clearly stated?
- Artefact correction, pre-rest period, duration of measurement at least 2 minutes, posture during measurement, specification frequency bands, measurement of breathing rate, 5-6 stated (2)
- Artefact correction, pre-rest period, duration of measurement at least 2 minutes, posture during measurement, specification frequency bands, measurement of breathing rate, 3-4 stated (1)
- Artefact correction, pre-rest period, duration of measurement at least 2 minutes, posture during measurement specification frequency bands, measurement of breathing rate, 1-2 or none stated (0)

7- Is outcome HRV clearly described and presented?
- Appropriate units (absolute units or normalized units) and measures of dispersion stated (2)
- Only appropriate units but no measures of dispersion stated (1)
- Outcome not clearly stated (0)

APPROPRIATE CONTROL FOR CONFOUNDING

8- Are potential confounders assessed[a]?
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 5-7 stated (2)
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 3-4 stated (1)
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 1-2 or none stated (0)

9- Are the analyses adjusted for potential confounders[b]?
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 5-7 stated (2)
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 3-4 stated (1)
- Age, gender, body mass index, smoking, depression, anxiety, medication use, 1-2 or none stated (0)

[a] *In case of exclusion at item 3, consider confounder as assessed.*

[b] *In case of exclusion at item 3 or no significant difference between cases and controls at item 8 consider confounder as adjusted for.*

the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996). We have chosen six important conditions for a reliable and reproducible HRV measurement that should be reported: conducting of artefact correction, presence of pre-rest period, duration of measurement at least two minutes, definition of frequency bands, posture during measurement, and measurement of breathing rate. The rationales for the first four conditions have been captured by the Task Force paper. Position of participants during the measurement influences HRV (Sipinkova et al. 1997). Monitoring and adjustment of spontaneous breathing rate during HRV measurements has been recommended (Berntson et al. 1997), but it has also been demonstrated that PNS activity is not dependent on spontaneous respiration frequency under baseline conditions (Denver et al. 2007). As the discussion is not closed yet, we recommend monitoring breathing rate and performing a secondary analysis with adjustment for breathing rate. A uniform outcome measure in HRV studies in FSD is lacking, especially in spectral analysis. The Task Force paper recommends reporting absolute values or normalized units, but suggests stating also absolute values in the latter to provide a complete description of the distribution of power in the spectral components. Normalized HRV-LF and HRV-HF units are frequently presented as two different variables. However, the mathematical procedure underlying calculation of normalized units dictates that HRV-HF normalized units and HRV-LF normalized units are equivalent carriers of information about ANS function (Burr 2007). The time-domain measure RMSSD has to be reported in ms.

Key domain 3 (items 8-9): Appropriate control for confounding
There are a number of potential confounders in the relation between HRV and FSD. Gender and age (Britton et al. 2007), smoking (Hayano et al. 1990), body mass index (Britton et al. 2007), anxiety (Cohen & Benjamin 2006), depression (Rottenberg 2007), and frequency of exercise (Britton et al. 2007) may be responsible for variance in ANS function. Gender (Kroenke & Spitzer 1998a), age (Escobar et al. 1987), smoking (John et al. 2004), body mass index (Neumann et al. 2008), anxiety, and depression (Henningsen et al. 2003) are also associated with FSD.

We have included those variables in our quality tool. Based on those nine items in three key domains, a judgment of quality was made by two independent reviewers (L.M.T and I.C.K.). We tested interrater reliability by calculating the percent of agreement between the reviewers. The maximum attainable quality score for a study was 18 points.

Statistical analysis
Meta-analyses were carried out in STATA 10.0 (StataCorp, College Station, TX) using the user-contributed command METAN (Bradburn et al. 1998). Each study's SMD was weighted by its inverse variance and an accompanying 95% confidence

interval (95% CI) was calculated. The random effects model, allowing for between-study variation of effect sizes, was considered more plausible a priori because of previously observed differences in methodological study quality (Tak & Rosmalen 2007). Therefore, a random effects model was fitted and presented in the forest plot. We then performed subgroup analyses to examine whether the summary estimate (weighted mean SMD) was moderated by the type of FSD. The *Q* test was performed to examine whether there was more heterogeneity in the effect sizes than could be expected from chance alone (DerSimonian & Laird 1986b). Additionally, we calculated the I^2 statistic, expressing the percentage of total variation that can be attributed to heterogeneity rather than chance. Publication bias was visually evaluated by a funnel plot and quantified by Egger's test (Egger et al. 1997a). In case of suspected bias, we performed a trim and fill procedure as sensitivity analysis (Duval & Tweedie 2000a). The trim and fill method allows estimation of an adjusted meta-analysis estimate in the presence of publication bias. Within the scope of our systematic review of methodological quality, we calculated correlation coefficients for the association between study quality and number of participants (Spearman's rho) and year of publication (Pearson's *r*). To test whether scores on items 5 and 6 of the quality tool improved after publication of the Task Force paper (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996), we performed a chi squared test. We used ANOVA to test whether quality differed between the three FSD. To test whether methodological study quality predicted author conclusions, we calculated an odds ratio (OR) with study conclusion (0 = no difference, 1 = lower PNS activity in FSD patients) as outcome variable. Finally, we determined whether methodological study quality was correlated with the SMD as calculated in the meta-analysis. All *p*-values are two-tailed and values less than 0.05 were considered statistically significant.

RESULTS

Search results and description of studies

The search strategy identified a total of 271 references. Another six references were retrieved from searching reference lists. The consultation of experts about (un)published data did not lead to additional inclusion of articles. After applying inclusion criteria, 54 references were left for full text reading. Applying our exclusion criteria led to a further exclusion of 31 articles, with as most frequent reason double publication of data on the same subjects in different articles. In summary, 23 articles met eligibility criteria for our systematic review (see Appendix).

Table 2 lists the 23 available studies, which together included 533 FSD cases. Five studies reported on CFS, seven studies on FM, and 11 studies on IBS. The average number of cases was 24 (range 8 – 70), whereas the average number of controls was 20 (range 8 – 38). The mean average age of cases was 40 (range 29 –55) years; the mean average age of controls was 39 (range 28 – 53) years.

Authors of five out of 23 studies (22%) reported significantly lower PNS activity in FSD patients compared to healthy controls, while the other 18 studies (78%) did not report significant differences. None of the studies reported significantly higher PNS activity in FSD patients compared to controls.

Meta-analysis of PNS activity in patients with FSD compared to healthy controls
Of the total of 23 studies, 14 did report an outcome in log-transformed or absolute HRV values, rendering them suitable for inclusion in a meta-analysis (raw outcome data in Table 2). This meta-analysis revealed that across those 14 studies, PNS activity was significantly lower in FSD patients compared to healthy controls (mean weighted SMD = -0.32, 95% CI -0.63 to -0.01, z = -2.03, $p = 0.04$). There was significant heterogeneity in effect sizes (Q test $\chi^2 = 37.07$, $p<0.001$; $I^2 = 65\%$). See Figure 1 for a forest plot.

Next, we performed subgroup analyses to assess whether the type of FSD moderated the mean weighted SMD of PNS activity. The summary estimate for PNS activity in CFS ($N = 3$) was not significant (mean weighted SMD -0.12, 95% CI -0.48 to 0.25, $z = 0.64$, $p = 0.52$). There were no indications for significant heterogeneity in effect sizes (Q test $\chi^2 = 4.08$, $p = 0.25$; $I^2 = 27\%$). The summary estimate for PNS activity in FM ($N = 6$) was not significant (mean weighted SMD -0.55, 95% CI -1.33 to 0.23, $z = 1.37$, $p = 0.17$). Indications for substantial heterogeneity in effect sizes were present (Q test $\chi^2 = 25.99$, $p<0.001$; $I^2 = 85\%$). The summary estimate for PNS activity in IBS ($N = 5$) was not significant (mean weighted SMD -0.21, 95% CI -0.58 to 0.16, $z = 1.10$, $p = 0.17$). There were no indications for significant heterogeneity in effect sizes (Q test $\chi^2 = 5.05$, $p = 0.28$; $I^2 = 21\%$). Those subgroup analyses indicate that, although the mean weighted SMD is the largest for FM, PNS activity findings in FSD are not clearly moderated by type of FSD.

Table 2. *Characteristics of 23 selected studies for systematic critical review.*

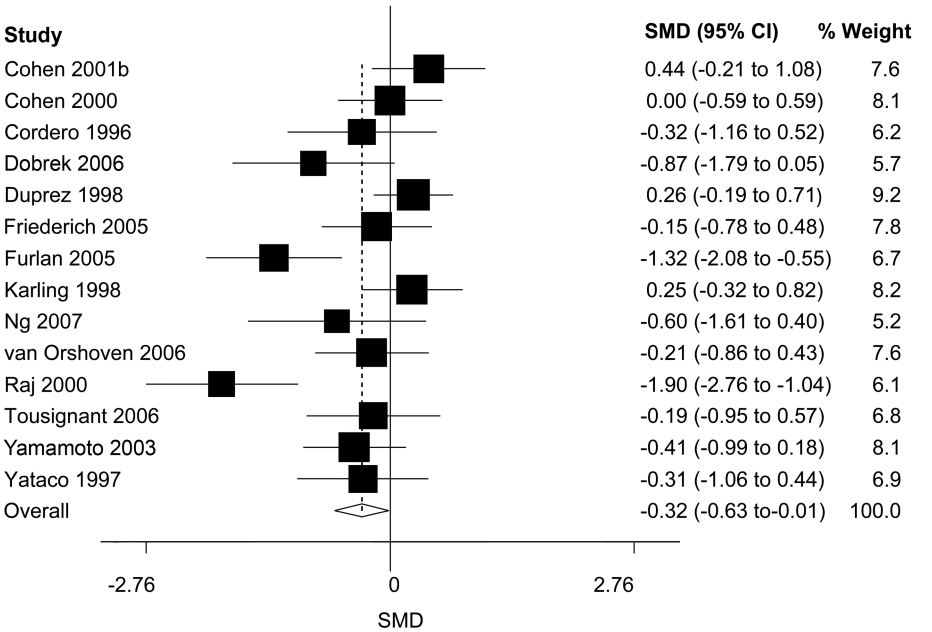
Author	FSD	*N*	Quality[a]	PNS activity in FSD[b]			
				Author conclusions	*Log transformed values*	*Absolute values*	*Normalized units*
Adeyemi et al. 1999	IBS	Cases 35 Controls 18	6	No difference	Med 2.4 (0.2 - 30.3) bpm^2/Hz Med 1.8 (0.1 - 10.1) bpm^2/Hz	na	na
De Becker et al. 1998	CFS	Cases 21 Controls 13	10	No difference	na	na	0.423 SD 0.164 nu 0.358 SD 0.170 nu
Cohen et al. 2001b	FM	Cases 19 Controls 19	13	Lower	4.41 SD 0.87 log ms^2 3.99 SD 1.05 log ms^2	na	8.2 SD 3.6 nu 22.6 SD 12.5 nu
Cohen et al. 2000	FM	Cases 22 Controls 22	15	Lower	2.83 SD 1.88 log ms^2 2.83 SD 2.78 log ms^2	na	13 SD 6 nu 42 SD 24 nu
Cordero et al. 1996	CFS	Cases 11 Controls 11	7	No difference	2.75 SD 1.39 log ms^2 3.09 SD 0.56 log ms^2	na	na
Cordova 2005	FM	Cases 25 Controls 29	5	No difference	na	na	0.12 nu 0.09 nu
Dobrek et al. 2006	IBS	Cases 10 Controls 10	8	Lower	na	271.5 SD 277.2 ms^2 854.6 SD 1364.9 ms^2	na
Duprez et al. 1998	CFS	Cases 38 Controls 38	9	No difference	na	1144 SD 1601 ms^2 971 SD 1632 ms^2	na
Elsenbruch & Orr 2001	IBS	Cases 24 Controls 20	11	No difference	na	na	0.43 SD 0.68 dB 0.37 SD 0.31 dB
Friederich et al. 2005	FM	Cases 28 Controls 15	9	No difference	4.2 SD 0.66 log ms^2 4.3 SD 0.66 log ms^2	na	28.2 SD 18.5 nu 34.0 SD 22.7 nu
Furlan et al. 2005	FM	Cases 16 Controls 16	11	Lower	na	198 SD 204 ms^2 939 SD 1460 ms^2	25.4 SD 15.2 nu 46.0 SD 17.6 nu
Karling et al. 1998	IBS	Cases 18 Controls 36	7	No difference	3.8 SD 0.4 mHz^2 3.7 SD 0.4 mHz^2	na	na
Martinez-Lavin et al. 1997	FM	Cases 19 Controls 19	12	No difference	na	na	0.242 SD 0.154 nu 0.289 SD 0.176 nu
Ng et al. 2007	IBS	Cases 8 Controls 8	9	No difference	na	230 SD 205 ms^2 441 SD 512 ms^2	na
Orr et al. 2000	IBS	Cases 14 Controls 15	9	No difference	na	na	20 SD 12 % power 26 SD 16 % power
van Orshoven et al. 2006	IBS	Cases 18 Controls 19	9	No difference	na	1017 SD 1536 ms^2 1030 SD 1135 ms^2	na
Raj et al. 2000	FM	Cases 17 Controls 14	9	Lower	5.45 SD 0.83 ln ms^2 7.03 SD 0.83 ln ms^2	na	na
Robert et al. 2004	IBS	Cases 70 Controls 21	10	No difference	na	na	Not provided
Thompson et al. 2002	IBS	Cases 33 Controls 21	11	No difference	na	na	49 SD 26 % power 50 SD 31 % power

Table 2 (continued). *Characteristics of 23 selected studies for systematic critical review.*

Author	FSD	*N*	Quality[a]	PNS activity in FSD[b]			
				Author conclusions	*Log transformed values*	*Absolute values*	*Normalized units*
Tousignant-Laflamme et al. 2006	IBS	Cases 14 Controls 13	11	No difference	na	1181 SD 3278 ms^2 698 SD 669 ms^2	34 SD 19 nu 37 SD 18 nu
Waring et al. 2004	IBS	Cases 30 Controls 30	12	No difference	na	na	47 SD 16 nu 46 SD 16 nu
Yamamoto et al. 2003	CFS	Cases 24 Controls 22	9	No difference	9.07 SD 7.23 sqrt ms^2 11.80 SD 6.14 sqrt ms^2	130 SD 222 ms^2 170 SD 159 ms^2	na
Yataco et al. 1997	CFS	Cases 19 Controls 11	7	No difference	3.4 SD 2.4 no unit given 4.2 SD 2.9 no unit given	na	na

Abbreviations: CFS = chronic fatigue syndrome, FM = fibromyalgia, FSD = functional somatic disorder, HF = high frequency; HRV = heart rate variability; IBS = irritable bowel syndrome, na = not applicable, nu = normalized units, PNS = parasympathetic nervous system. Means and SD are provided, unless otherwise stated. Med = median (range). [a] *Quality as measured with quality tool (Table 1).* [b] *All outcomes are HRV-HF values, as none of the selected studies reported the root mean square of differences between successive interbeat intervals (RMSSD) only.*

Figure 1. *Forest plot demonstrating standardized mean differences (SMD) of the individual studies and the mean weighted effect size of the association between functional somatic disorders and decreased parasympathetic activity as measured with heart rate variability. A random effects model is used.*



Finally, we tested whether the significantly lower PNS activity in FSD patients compared to healthy controls could be explained by publication bias. By visual inspection, asymmetry of the funnel plot was suspected (Figure 2A). Accordingly, Egger's test rejected the null hypothesis that there was no funnel plot asymmetry (p = 0.01). After performing the trim and fill procedure, 5 studies were filled (Figure 2B). The sensitivity analysis using the trimmed and filled weighted SMD of 19 studies indicated that there is no significant association between FSD and PNS activity anymore when taking potentially missing studies into account (mean weighted SMD = 0.01, 95% CI -0.34 to 0.36, z = 0.07, p = 0.95).

Does methodological quality predict the finding of altered PNS activity?
Interrater reliability of quality scoring was 95%; most differences occurred due to reader error or ambiguity in the paper. Methodological quality score was normally distributed (skewness = 0.24, kurtosis = 0.26). Mean quality score was 9.5 (SD 2.3) points and the maximum obtained quality score was 15 points. Mean quality was not significantly different for any of the three FSD ($F_{(2,20)}$ = 1.35, p = 0.28). Furthermore, study quality was not significantly associated with the number of cases in a study (Spearman's rho 0.12, p = 0.60).

Figure 2A. *Funnel plot (N = 14) showing the correlation between the standardized mean difference (SMD) and its standard error (SEM) with pseudo 95% confidence limits.*

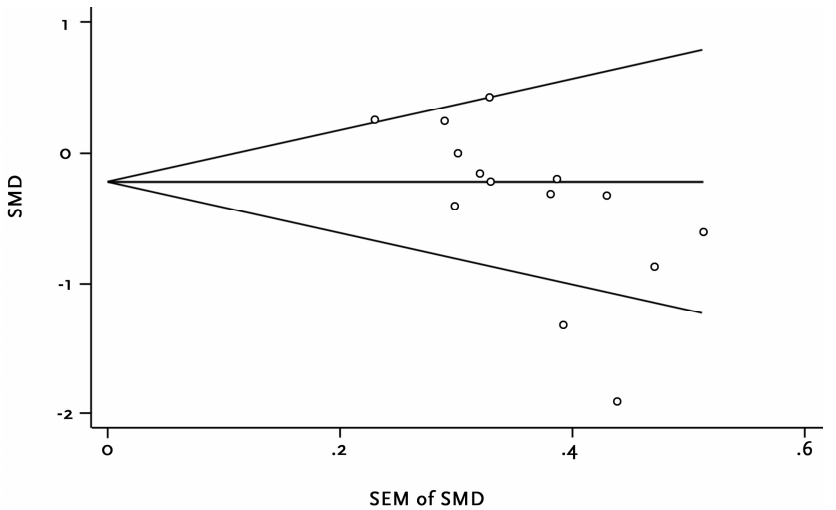
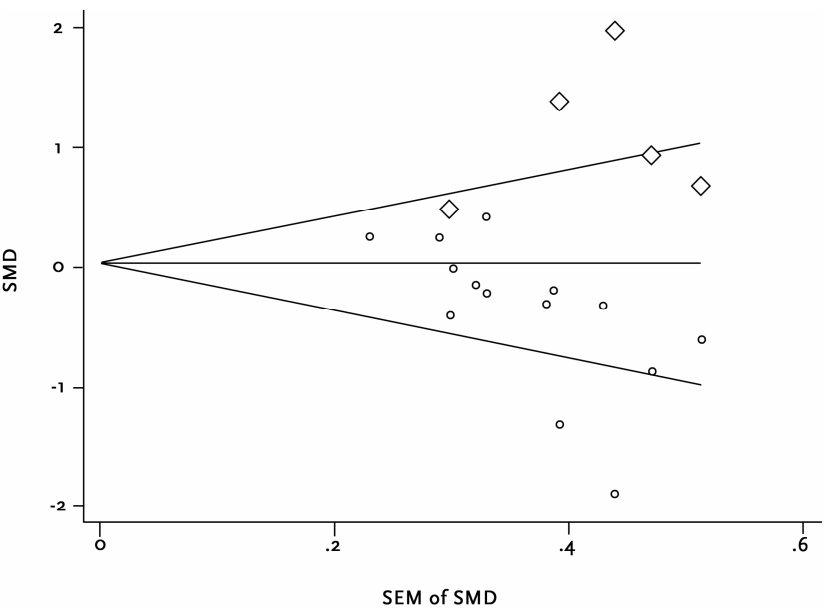


Figure 2B. *Trimmed and filled funnel plot (N = 19) showing the correlation between the standardized mean difference (SMD) and its standard error (SEM) with pseudo 95% confidence limits. Squares represent the studies that have been filled.*



To explore how studies performed on different key domains of methodological quality, first, we examined whether participants were appropriately selected. In contrast with the selection of cases, which was generally in conformity with international criteria with involvement of a physician, selection of controls was poor because they were frequently recruited from hospital personnel. Exclusion of somatic co-morbidity was usually stated, whereas psychiatric co-morbidity and medication use were not always mentioned. Disease characteristics are described adequately only in a minority of the studies. Concerning the second key domain, appropriate measurement of HRV, only one study explicitly stated that the assessor of HRV was blinded for disease status. Overall, studies scored moderate to good on the report of methods for assessment of HRV. Often, outcome measures were not clearly presented in absolute values or normalized units. Assessment and, when necessary, appropriate adjustment for potential confounders as measured in the third key domain, leaves room for improvement. None of the studies adjusted for the possible confounder depression, for instance. See Table 3 for an overview of methodological study quality.

Table 3. *Overview of quality scores per item of the quality tool.*

Quality item	0 points *N* (%)	1 point *N* (%)	2 points *N* (%)
APPROPRIATE SELECTION OF PARTICIPANTS			
(1) Disease reliably assessed and validated?	0 (0)	7 (30)	16 (70)
(2) Controls recruited from same population as cases?	5 (22)	10 (43)	8 (35)
(3) Population defined with in- and exclusion criteria?	0 (0)	18 (78)	5 (22)
(4) Disease characteristics presented?	10 (44)	9 (39)	4 (17)
APPROPRIATE QUANTIFICATION OF HRV			
(5) Assessor of HRV blind for disease status?	22 (96)	-	1 (4)
(6) Methods for assessment of HRV clearly stated?	1 (4)	10 (44)	12 (52)
(7) Outcome HRV clearly described and presented?	9 (39)	3 (13)	11 (48)
APPROPRIATE CONTROL FOR CONFOUNDING			
(8) Potential confounders assessed?	2 (9)	13 (57)	8 (35)
(9) Statistically adjusted for confounders?	5 (22)	17 (74)	1 (4)

Abbreviation: HRV = heart rate variability.

Next, we aimed to determine whether quality has improved over time. Year of publication and quality score were not significantly correlated (Pearson's $r = 0.07$, $p = 0.74$). Furthermore, we found no evidence that methods for assessment or reporting appropriate outcome measures (item 5 and 6) have improved after publication of the Task Force paper ($\chi^2 = 2.94$, $df = 4$, $p = 0.57$). Study quality was a borderline significant predictor for study conclusions ($N = 23$, OR 1.60, 95% CI

0.93 to 2.75, $z = 2.83$, $df = 1$, $p = 0.09$), suggesting that higher quality studies have a higher chance of concluding that PNS activity in FSD patients is significantly lower than in healthy controls. In the subgroup of studies included in the meta-analysis, total quality score and SMD were not significantly correlated ($N = 14$, $r = 0.16$, $p = 0.59$).

DISCUSSION

The aims of the current study were to assess whether FSD are characterized by alterations in PNS activity and to examine whether this association was moderated by the type of FSD or influenced by methodological study quality. Meta-analysis of 14 published studies revealed lower PNS activity in FSD patients compared to controls that was not significantly moderated by type of FSD (i.e., CFS, FM, or IBS). The reliability of all summary estimates in this meta-analysis is, however, significantly limited by unexplained heterogeneity in effect sizes and potential publication bias. The systematic review of overall methodological quality of HRV studies in FSD demonstrates that there is substantial room for improvement, especially in selection of control subjects, blinding of researchers performing HRV measurements, report of adequate HRV outcomes, and assessment and adjustment for potential confounders. Although higher methodological study quality tended to predict author conclusions on decreased PNS activity in FSD, an association between methodological study quality and effect size was not found in the meta-analysis.

When extracting outcome measures of the studies, we concluded that the variety in outcome measures between studies was large, and even worse, reported outcomes were sometimes physiologically or mathematically impossible. We were doubtful about overall statistical accuracy of studies performing parametric tests (*t*-test, for instance) on probably non-normally distributed HRV measures in small sample sizes (Duprez et al. 1998, Furlan et al. 2005, Dobrek et al. 2006, Ng et al. 2007).

We expected a large heterogeneity in effect sizes across the different studies, due to aforementioned data irregularities and differences in methodological quality. This was confirmed by the presence of substantial statistical heterogeneity in effect sizes as revealed in the meta-analysis. A solution for heterogeneity is finding moderators that explain this heterogeneity (Hardy & Thompson 1998), such as the type of FSD, selection of physically inactive controls, exclusion of subjects using cardio-active medication, the type of PNS activity outcome measure, or methodological quality. However, it has been discouraged to perform subgroup analyses when the total number of included studies in the meta-analysis is small. In this case, the insufficient degrees of freedom would inflate

the risk for spurious findings (Higgins et al. 2002b, Babyak 2004). Therefore, we decided to restrict subgroup analysis to the three different types of FSD.

A general problem affecting the validity of meta-analyses and systematic reviews is publication bias. In the present meta-analysis, the asymmetric funnel plot and formal test suggested the presence of significant publication bias. After correction of funnel plot asymmetry, the statistically significant mean weighted SMD disappeared. These results suggest that some studies contrasting with the prevailing beliefs on lowered PNS function in FSD have not been published. It is important to note, however, that publication bias may not be the only cause of funnel plot asymmetry. Potential other sources of asymmetry in funnel plots that could be particularly applicable to our meta-analysis are data irregularities due to poor methodological design of smaller studies and inadequate analyses (Egger et al. 1997a). It should be noted, however, that we did not find evidence that smaller studies had lower quality scores. When there is large between-study heterogeneity in effect sizes, trim and fill models may erroneously adjust for non-existing publication bias (Peters et al. 2007). If the unadjusted and adjusted summary estimates differ considerably, as in our meta-analysis, it has been advised to consider the findings with caution whatever the cause of funnel plot asymmetry.

Some limitations of our meta-analysis and systematic review should be noted. First, although face validity of our quality tool may seem adequate, it has not been proven that quality scoring in experimental research is really valid (Herbison et al. 2006). Furthermore, due to reasons of comprehensibility of our quality tool (Sanderson et al. 2007), we could not list all conditions that are important for a reliable and reproducible HRV measurement. Time of day of HRV measurement (van Eekelen et al. 2004a), as well as restrictions in behaviors shortly prior to HRV measurement, such as smoking (Karakaya et al. 2007), intake of alcohol (Koskinen et al. 1994), intake of caffeine (Sondermeijer et al. 2002), water ingestion (Routledge et al. 2002), and ingestion of a meal (Lu et al. 1999) may influence the results. We recommend restricting from those behaviors two hours before the HRV measurement; longer restriction may cause withdrawal effects that also influence HRV, which is established for smoking (Yotsukura et al. 1998) but may also account for the other substances. Furthermore, only the most important confounders according to the experts' opinion are listed in our quality tool. Other potentially important confounders are a sedentary lifestyle instead of habitually exercising (Houtveen & van Doornen 2007b, Britton et al. 2007) and daily alcohol use (Thayer et al. 2006, Hasin & Katz 2007). It should, however, be realized that none of the included studies is large enough to adjust for all confounders currently included in the quality tool. Despite of its limitations, quality scoring gives us the possibility to provide criteria which can be applied to future HRV studies in subjects suffering from FSD, and after slight adaptations also to studies measuring HRV in subjects suffering from other mental disorders. Second, we have only stated recommendations for baseline HRV measurements

at rest. We did not review the dynamic aspect - reactivity or homeostatic capacity - of the ANS, while dysfunction may be more pronounced after provocation tests. Nevertheless, a review with recommendations or gold standard for measuring HRV after challenge tests is also urgently needed. In an earlier narrative review, it seems that FSD are characterized by lower PNS activity and diminished reactivity after challenge tests (Tak & Rosmalen 2007).

We have some recommendations concerning how to proceed in the field of HRV studies in FSD. Selection bias can be diminished by careful selection of participants or by nested case-control studies in population cohorts. For the purpose of this review, we have only selected studies on one of the main three FSD, because selection of cases based one FSD exclusively has been usual practice when studying HRV in FSD. Bearing in mind that significant cross-over co-morbidity among FSD exists (Aaron & Buchwald 2001), it may be a tenable approach for future HRV studies to select patients that experience functional somatic symptoms across different FSD (Houtveen & van Doornen 2007b) or to include cases of the three different FSD in one study. It remains recommendable to test whether psychophysiological variables, such as PNS activity, are differentially related to specific symptom clusters or are a common underlying factor. Finally, future studies should preferably have a prospective design to examine whether, if present, dysfunction of the ANS is an etiological factor in FSD, a consequence of FSD, or that ANS dysfunction and FSD covary due to a third variable that affects both.

Consensus about reporting HRV outcomes would mean a major improvement in the field. Although absolute spectral values of HRV in ms^2 are known to have a skewed distribution, which is usually normalized after natural logarithmic transformation, interpretation is problematic since normative data are lacking and it appears difficult to compare data from different laboratories. The growing body of studies using HRV, with over 500 peer-reviewed publications annually (Sandercock 2007), illustrates the importance of consensus toward a normative range for HRV indices. A clear need for consensus about in which situation either absolute values or normalized units are preferable is illustrated by the following example. Authors of a study in FM patients concluded that PNS activity was decreased based on their outcome reported in normalized units (Cohen et al. 2001b). However, the effect size of that specific study in our meta-analysis, based on log transformed values, pointed to the opposite direction. In addition, in an attempt to study sympathetic nervous system activity, previous studies have often reported power in the low frequency band (HRV-LF), defined at 0.04 - 0.15 Hz. However, the physiological basis is not well understood and HRV-LF certainly does not simply reflect sympathetic activity (Berntson et al. 1997), and should not be reported as such. Furthermore, future studies should be large enough to allow adjustment for at least seven important confounders. It is worrisome that studies which are methodologically severely flawed or which present results that are

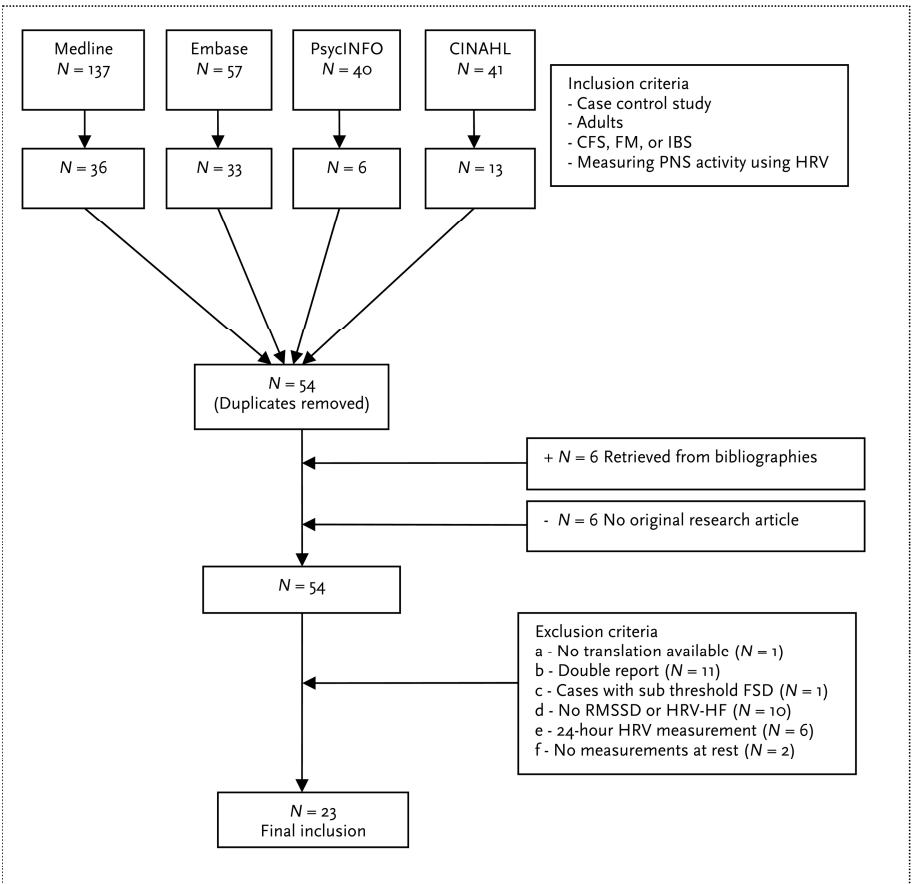
mathematically impossible have been published. Prevention of this is not only a task for authors, but also for peer-reviewers and editors. Furthermore, to ensure that journal articles are a representative sample of research conducted in this field, authors should be stimulated to submit negative or counterintuitive findings to scientific journals, while editors should be encouraged to publish them.

We conclude that current available evidence is not adequate to firmly reject or accept a role of ANS dysfunction in FSD and it would therefore be misleading to provide a definitive summary estimate. Nevertheless, apart from providing a summary estimate, the importance of this meta-analysis also lies in the identification of heterogeneity of effect sizes and the construct of guidelines to conduct better research in the future. Current overall methodological quality of studies examining HRV in FSD should not be as good as it gets, but can be substantially improved. We have given suggestions to help overcome the limitations that we have identified. Future research adhering to quality recommendations for HRV studies in FSD patients may hopefully provide a more definite answer on the question whether ANS dysfunction is involved in the etiology of FSD.

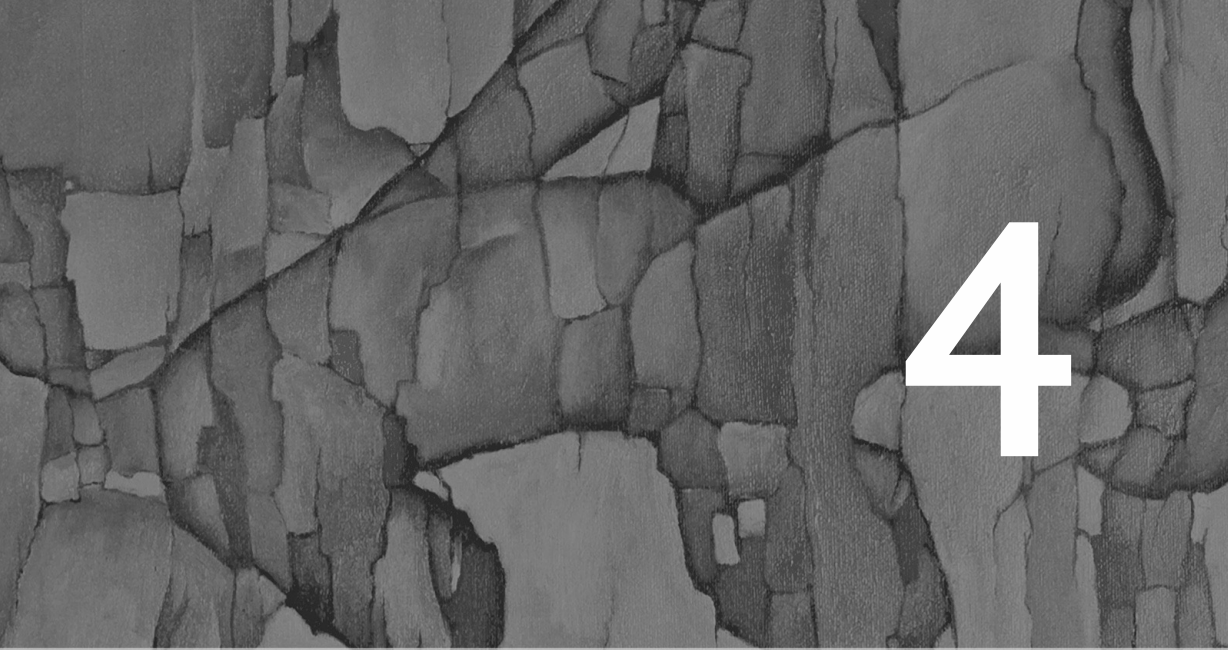
ACKNOWLEDGEMENTS

We thank the experts (dr. S. Elsenbruch, prof. W. Orr, prof. M.M. Heitkemper, dr. R.L. Burr, dr. K. Cain, dr. M. Martinez-Lavin, and dr. K. Yoshiuchi) for their contribution to the quality tool and dr. ir. A.M. van Roon for his methodological advices regarding heart rate variability measurements

Appendix. *Results of systematic search after applying in- and exclusion criteria.*



Abbreviations: CFS = chronic fatigue syndrome, FM = fibromyalgia, HF = high frequency, HRV = heart rate variability, IBS = irritable bowel syndrome, PNS = parasympathetic nervous system, RMSSD = root mean square of successive interbeat interval differences.



4

CHAPTER 4

Meta-analysis and meta-regression of hypothalamic-pituitary-adrenal axis activity in functional somatic disorders

LM Tak, AJ Cleare, J Ormel, A Manoharan, IC Kok, S Wessely, and JGM Rosmalen
Submitted

ABSTRACT

Dysfunction of the hypothalamic-pituitary-adrenal (HPA) axis is the most investigated biological risk factor in functional somatic disorders (FSD), such as chronic fatigue syndrome (CFS), fibromyalgia (FM), and irritable bowel syndrome (IBS). However, findings over the past 20 years are conflicting. The objective of this study was to assess whether there is an association between basal hypocortisolism in FSD and to identify potential moderators (type of FSD, percentage females, medication use, co-morbid depressive disorder, and physical inactivity) of this association. A standardized mean difference (SMD) between FSD cases and controls of basal cortisol levels in either saliva, serum, or urine was calculated. Random effects models were applied. Meta-analysis on eighty-five studies revealed that although basal cortisol levels were generally lower in FSD subjects compared to controls, this association did not reach statistical significance (SMD -0.07, 95% CI -0.17 to 0.04, p = 0.241). However, when the three FSD were assessed separately, statistically significant hypocortisolism was observed in CFS subjects compared to controls (SMD -0.14, 95% CI -0.28 to 0.00, p = 0.047), but not in FM or IBS. When all potential moderators were entered into a meta-regression analysis, only type of FSD and female gender were significant independent predictors of hypocortisolism. We did not find evidence to consider all FSD as hypocortisolemic disorders, as significant reduction in cortisol compared to healthy controls was only found in CFS, in females with FM, but not in IBS.

INTRODUCTION

Functional somatic disorders (FSD) are syndromes of related physical complaints without known underlying conventional organic pathology. The main three disorders are chronic fatigue syndrome (CFS), fibromyalgia (FM), and irritable bowel syndrome (IBS). Shared factors might underlie general susceptibility for development of any FSD, whereas FSD-specific factors might shape their final manifestation (Aggarwal et al. 2006, Kato et al. 2009).

HPA axis dysfunction, the most widely investigated biological factor in the etiology of FSD, is one potential shared factor, as alterations in this stress responsive system have been reported for all main FSD (Tak & Rosmalen 2007). A potential etiological link between the HPA axis and FSD emerges from the potential of HPA axis underactivity to increase symptoms through mechanisms such as increasing pain perception and causing fatigue (Lariviere & Melzack 2000, Heim et al. 2000a, Fries et al. 2005, Fabian et al. 2009). Some even already refer to FSD as hypocortisolemic disorders (Fries 2008). However, narrative reviews conclude that findings on cortisol levels in CFS, FM, and IBS subjects compared to healthy controls are inconsistent: as well as mild hypocortisolism, normal or increased cortisol levels have also been reported (Mayer et al. 2001, Geenen et al. 2002, Cleare 2003, Tak & Rosmalen 2007). It remains to be elucidated why the presence of HPA axis alterations varies both within and among FSD. Moreover, if FSD are really characterized by HPA axis dysfunction, its position in the etiological pathway, if causally linked at all, is still elusive. The idea of HPA axis dysfunction as a mediator between psychosocial stress and FSD has often been advanced, which is supported by the observation that retrospective psychosocial stress has been consistently associated with FSD (Barsky & Borus 1999, van Houdenhove et al. 2005, Aggarwal et al. 2006, Deary et al. 2007) and has the capacity to induce hypocortisolism in the long-term (Ehlert et al. 2001, Miller et al. 2007, McEwen 2007). However, HPA axis alterations could also be a consequence of factors such as concurrent stress, sleep disturbances, alcohol use, smoking, obesity, medication use, co-morbid depressive disorder, or physical inactivity (Geenen et al. 2002, Cleare 2003). Twenty years of research has given rise to conflicting findings, suggesting a need for a systematic meta-analysis of previous research.

The primary purpose of this meta-analysis is to quantify the association between HPA axis function and FSD. We hypothesize that FSD are characterized by hypocortisolism. Additionally, we hypothesize that HPA axis dysfunction is a shared factor for all FSD and hypocortisolism therefore manifests irrespective of the diagnostic label of CFS, FM, or IBS. A second goal is to identify potential moderators of the association between HPA axis function and FSD, including gender, medication use, co-morbid depressive disorder, and physical inactivity.

METHODS

Search strategy

Relevant articles were identified by searching the databases of Medline, Embase, and PsycINFO (January 1960 - November 2009). A search string was formulated for searching Medline. The first component consisted of chronic fatigue syndrome, fibromyalgia, irritable bowel syndrome, and synonyms. The second component consisted of the terms hypothalamic-pituitary-adrenal axis, cortisol, and synonyms. Searching Embase and PsycINFO, terms included in our search were adapted according to the thesaurus of the respective database and explosion of the search terms was applied. Reference lists of original articles and related reviews were hand searched for additional citations. The search was conducted without language restrictions.

Screening and selection procedure

Title and abstract of the articles were screened on three inclusion criteria (case-control studies; cases are adults with CFS, FM, or IBS; measurement of HPA axis) by two independent reviewers. Discrepancies were resolved by consensus. Full text articles were acquired and screened on the following exclusion criteria: a) duplicate reports of the same subjects, b) HPA axis not measured under baseline conditions, c) no HPA axis data of healthy controls presented, d) no extractable data or statements on HPA axis function, or e) not original research.

Data extraction

From every included paper, name of first author, year of publication, type of FSD, number and age of participants, potential moderators (see 2.4) and baseline cortisol levels in either serum, saliva, or urine were extracted. Previous meta-analyses on HPA axis activity in psychobiological studies lend support to the comparability of cortisol assays in either saliva, blood, or urine (Meewisse et al. 2007, Michaud et al. 2008). When multiple measurements during the day were performed, the measurement closest to 0800h was selected as the morning measurement and the measurement closest to 1600h was selected as the afternoon measurement. For the primary meta-analysis combining various types of assays, the hierarchy of the selected measurement was as follows: morning samples were preferred above afternoon samples, saliva was preferred above serum and serum was preferred above 24-hour urinary free cortisol (24-h UFC) (Kirschbaum & Hellhammer 1994, Meewisse et al. 2007). If data were not extractable, i.e., no appropriate units and measures of dispersion or test statistics were provided, authors were contacted and asked for additional information. For ten studies, authors supplied additional outcome information that was not presented in the original publication (Elsenbruch et al. 2002, Zarkovic et al. 2003, Elsenbruch et al. 2004, Posserud et al. 2004, Bohmelt et al. 2005, Meeus et al. 2008, Klingmann et al. 2008, Nater et al. 2008a, Nater et al. 2008b, Chang et al. 2009). To allow pooling across studies that used different types of HPA axis

measurement, we calculated a standardized mean difference (SMD) of basal cortisol levels, which is an effect size measure based on the difference between mean values of FSD subject and healthy control groups divided by the pooled standard deviation (also referred to as Cohen's *d*) for each study (Rosenthal & DiMatteo 2001). All data extraction was done by two independent reviewers. In case of missing data, conservative effect sizes were estimated. The SMD of three studies only stating that cortisol levels in FSD patients were not significantly different from controls was therefore set at 0.00 (Yatham et al. 1995, Strickland et al. 1998, Malt et al. 2002). As introducing such conservative effect sizes may yield an underestimated summary effect size that underestimates the real value, the summary effect size was calculated without these three studies in the primary analysis but with those studies in a sensitivity analysis.

Coding of moderator variables

A second aim of this study was to assess sources of heterogeneity to explain differences between studies. In order to minimize the number of covariates investigated, we followed recommendations to select those justified through scientific rationale, and specify them in advance (Higgins & Thompson 2004, Freedland et al. 2009) Therefore we composed a priori a panel of variables that are most important potential sources of heterogeneity to be tested in moderator analyses, including type of FSD (either CFS, FM, or IBS), gender (based on median split of % females), medication use (exclusion or discontinuation of medication that affects the HPA axis, notably corticosteroids, oral contraceptives, estrogen replacement therapy, and antidepressants), co-morbid depressive disorder (exclusion of participants meeting diagnostic criteria for a depressive disorder) and physical activity (selection of matched physically inactive healthy controls). Some variables that might be theoretically important were rarely addressed in the original studies, including concurrent stress (N = 13), sleep disturbances (N = 13), and childhood trauma (N = 5), while others did not have enough variability between studies to enable the construction of different relevant subgroups, such as age, body mass index, smoking, and duration and severity of the FSD. Those variables could therefore not be tested in moderator analyses. Variables such as somatic co-morbidity and characteristics of HPA axis measurement may introduce noise in cortisol measurements but were not considered likely to essentially bias the results.

Quality assessment

To assess methodological study quality, we adapted a quality tool already developed for autonomic nervous system studies in FSD (Tak et al. 2009a) according to the specific characteristics of HPA axis studies (Table 1). Based on nine items in the three key domains selection of participants, measurement of HPA axis, assessment of confounders (see Appendix A for background), a judgment of quality was made by two independent reviewers. We tested interrater

reliability by calculating the kappa coefficient (Landis & Koch 1977). The maximum attainable quality score for a study was 18 points.

Table 1. *Quality tool to assess methodological quality of hypothalamic-pituitary-adrenal (HPA) axis function studies in functional somatic disorders.*

APPROPRIATE SELECTION OF PARTICIPANTS

1- Has the disease of the cases been reliably assessed and validated?
- According to international criteria by a physician (2)
- Not according to international criteria or assessor not clearly established (1)
- Only self report or not clearly stated (0)

2- Have all controls been recruited from the same population as the cases?
- Controls from same population as cases (2)
- Selected population, such as hospital staff or students (1)
- Not clearly stated (0)

3- Is the population defined with in- and exclusion criteria?
- Medication use, somatic morbidity, psychiatric morbidity, 3 stated (2)
- Medication use, somatic morbidity, psychiatric morbidity, 1-2 stated (1)
- None stated or not clearly stated (0)

4- Are disease characteristics presented (length and severity of functional somatic disorder)?
- Duration of disease and severity of disorder is stated (2)
- Only duration or only severity is stated (1)
- None stated (0)

APPROPRIATE QUANTIFICATION OF HPA AXIS FUNCTION

5- Is assessor of HPA axis blind for disease status?
- Yes (2)
- No or not clearly stated (0)

6- Have methods of HPA axis measurement been clearly stated*?
- Time of day, behaviour shortly prior to measurement, storage conditions, type of assay performed, repeated measurements, assessing compliance, 5-6 stated (2)
- Time of day, behaviour shortly prior to measurement, storage conditions, type of assay performed, repeated measurements, assessing compliance, 3-4 stated (1)
- Time of day, behaviour shortly prior to measurement, storage conditions, type of assay performed, repeated measurements, assessing compliance, 1-2 or none stated (0)

7- Is outcome HPA axis measurement clearly described and presented?
- Central tendency and measures of dispersion stated in appropriate units (2)
- Only central tendency but no measures of dispersion stated in appropriate units (1)
- Outcome not clearly stated (0)

APPROPRIATE CONTROL FOR CONFOUNDING

8- Are potential confounders assessed[a]?
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 5-7 stated (2)
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 3-4 stated (1)
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 1-2 or none stated (0)

9- Are the analyses adjusted for potential confounders[b]?
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 5-7 stated (2)
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 3-4 stated (1)
- Age, gender, body mass index, smoking, depression, medication, physical exercise, 1-2 or none stated (0)

[a] In case of exclusion at item 3, consider confounder as assessed.
[b] In case of exclusion at item 3 or no significant difference between cases and controls at item 8 consider confounder as adjusted for.

Statistical analyses

Meta-analyses were carried out by an independent statistician in STATA 10.0 (StataCorp, College Station, TX) using the user-contributed command METAN (Bradburn et al. 1998). Each study's SMD was weighted by its inverse variance and an accompanying 95% confidence interval (95% CI) was calculated. Given the previously found conflicting findings (Cleare 2003, Tak & Rosmalen 2007), the random effects model that allows for between-study variation of effect sizes was considered more plausible a priori. Therefore, random effects models were fitted and presented in the forest plot. The *Q* test was performed to examine whether there was more heterogeneity in the effect sizes than could be expected from chance alone (DerSimonian & Laird 1986). Additionally, we calculated the I^2 statistic, expressing the percentage of total variation that can be attributed to heterogeneity rather than chance. We then performed subgroup analyses to examine whether the summary effect size (weighted mean SMD) was moderated by an a priori defined set of variables. Meta-regression was performed with SMD as outcome variable and potentially moderating variables as predictor variables to assess their independent contributions. Regression coefficients and 95% confidence intervals (95% CI) were calculated. Sensitivity analyses taking into account missing values and study quality were performed. The influence of methodological study quality was studied by comparing the summary effect size in low versus high quality studies based on a median split. Publication bias was visually evaluated by a funnel plot and quantified by Egger's test (Egger et al. 1997a). We planned to perform a trim and fill procedure as an additional sensitivity analysis (Duval & Tweedie 2000a). All *p*-values less than 0.05 were considered statistically significant.

RESULTS

Search results and study characteristics

In total, we included 82 references, in which 85 case-control comparisons between FSD subjects and healthy controls were made (see Appendix B). Four of those 85 comparisons, did not provide means and standard deviations and were therefore not included in the primary meta-analysis.(Yatham et al. 1995, Strickland et al. 1998, Malt et al. 2002, Burr et al. 2009). Table 2 lists the 85 available comparisons, which together included 2148 FSD subjects and 1988 healthy controls. Forty studies reported on CFS (1010 cases; 1039 controls), 27 studies on FM (650 cases; 595 controls), and 18 studies on IBS (488 cases; 354 controls). Median number of FSD subjects was 20 (range 7 - 121); median number of controls was 17 (range 7 - 131). The average mean age of FSD subjects was 40 years (range 24 - 53); the average mean age of controls was 38 years (range 21 - 51). Median duration of the FSD was 62 months (range 18 - 212).

Table 2. *Characteristics of the included studies in the meta-analysis.*

Study	Type of FSD	*N* of cases	% F cases	Mean age cases	Mean duration (months)	*N* of con	% F con	Mean age con	Quality (points)	Cortisol saliva	Cortisol blood	Cortisol urine
Altemus et al. 2001	CFS	19	68	39,7	44	19	68	38,9	13		x	
Bohmelt et al. 2005	IBS	25	56	43,5	NR	24	56	38,4	12	x		
Burr et al.2009	IBS	30	100	30	NR	31	100	32	10		x	
Calis et al. 2004	FM	22	100	38,7	NR	15	100	36,5	11		x	
Catley et al. 2000	FM	21	86	47,9	47	22	86	46,6	15	x		
Chang et al. 2009	IBS	41	100	39,9	NR	25	100	33	12		x	
Cleare et al. 2001b	CFS	37	68	33,8	43	28	68	32,4	13		x	x
Cleare et al. 2001a	CFS	121	64	39,5	65	64	64	33,9	12			x
Cleare et al. 1995	CFS	10	40	36,5	NR	25	40	35	12		x	
Crofford et al. 2004	CFS	15	73	35	36	15	73	35,1	15		x	x
Crofford et al. 2004	FM	13	100	49,8	212	12	100	51	15		x	x
Crofford et al. 1994	FM	12	100	39,5	74	11	100	39,6	9		x	x
Demitrack et al. 1991	CFS	19	53	36,4	86	20	53	39,4	9		x	x
Dickhaus et al. 2003	IBS	15	60	39	NR	14	60	35	12		x	
Dinan et al. 2006	IBS	21	67	34,6	NR	21	67	30,2	11		x	
Dinan et al. 1997	CFS	14	80	38	NR	14	80	36,5	8		x	
Elsenbruch et al. 2006	IBS	17	76	42	NR	12	76	39	12		x	
Elsenbruch et al. 2004	IBS	14	100	47,7	169	14	100	40	13		x	
Elsenbruch et al. 2002	IBS	24	100	34	NR	17	100	36,4	12	x		
Elsenbruch et al. 2001	IBS	24	100	32,8	160	20	100	32,5	14	x		
Eriksson et al. 2008	IBS	80	91	NR	NR	21	91	NR	7		x	
Fukudo et al. 1998	IBS	10	50	23,8	NR	10	50	20,7	7		x	

Table 2 (continued). *Characteristics of the included studies in the meta-analysis.*

Study	Type of FSD	*N* of cases	% F cases	Mean age cases	Mean duration (months)	*N* of con	% F con	Mean age con	Quality (points)	Cortisol saliva	Cortisol blood	Cortisol urine
Gaab et al. 2002b	CFS	21	52	36	67	21	52	35,2	15	x		
Giske et al. 2008	FM	19	100	37	120	19	100	NR	13		x	
Griep et al. 1998	FM	40	90	43	128	14	90	38,1	15		x	x
Griep et al. 1993b	FM	18	100	38,3	114	18	100	36,8	17		x	
Griep 2000	FM	20	90	43,7	126	14	90	38,1	15		x	x
Griep 2000	CFS	12	75	43,4	196	14	75	38,1	15		x	x
Gur et al. 2004a	CFS	62	100	32,6	51	46	100	31,5	12		x	
Gur et al. 2004a	FM	68	100	31,4	48	46	100	31,5	12		x	
Gursel et al. 2001	FM	20	100	41,3	51	20	100	42,7	9		x	
Hamilos et al. 1998	CFS	7	86	43	NR	7	86	44,8	8		x	x
Hudson et al. 1999	CFS	20	60	37	NR	20	60	36	14		x	
Inder et al. 2005	CFS	12	NR	NR	NR	11	NR	NR	9		x	x
Izgi et al. 2005	CFS	20	70	37,6	NR	15	70	36,5	9		x	
Izquierdo-Alvarez et al. 2008	FM	47	100	53	NR	58	100	45,5	5			x
Jerjes et al. 2006b	CFS	28	50	34	25	27	50	32,6	12			x
Jerjes et al. 2005	CFS	15	53	35	32	20	53	33	17	x		
Kaufmann et al. 2008	FM	22	77	53,1	NR	22	77	51	10		x	
Kilkens et al. 2005	IBS	14	57	31,5	NR	14	57	32,5	11		x	
Kirnap et al. 2001	FM	16	81	37,3	NR	16	81	36,9	9		x	
Klerman et al. 2001	FM	10	100	39,7	NR	12	100	33,3	13		x	
Klingmann et al. 2008	FM	93	100	51,4	NR	100	100	44,4	6	x		
Light et al. 2009	FM	25	100	46,4	NR	31	100	40,6	9		x	

Table 2 (continued). *Characteristics of the included studies in the meta-analysis.*

Study	Type of FSD	*N* of cases	% F cases	Mean age cases	Mean duration (months)	*N* of con	% F con	Mean age con	Quality (points)	Cortisol saliva	Cortisol blood	Cortisol urine
Macedo et al. 2008	FM	27	85	49,4	NR	29	85	50,1	11	x	x	
MacHale et al. 1998	CFS	30	63	44,2	62	15	63	41,1	12		x	
Maes et al. 1998	FM	14	79	51	NR	17	79	41,8	8			x
Malt et al. 2002	FM	22	100	45	NR	13	100	43	7		x	
McLean et al. 2005	FM	20	80	43	NR	16	80	39	13	x		
Meeus et al. 2008	CFS	31	68	45	NR	31	68	44	7	x		
Moorkens et al. 2000	CFS	29	69	39,1	18	9	69	32,4	9		x	
Morriss et al. 2002	CFS	9	50	46	NR	9	50	45,2	15		x	
Nater et al. 2008b	CFS	24	79	49,6	151	36	79	49,9	14	x		
Nater et al. 2008a	CFS	75	77	43,9	90	110	77	44,8	17	x		
Ottenweller et al. 2001	CFS	17	100	34,4	NR	14	100	34,4	6		x	
Paiva et al. 2002	FM	20	100	44,6	NR	10	100	47	7		x	
Patacchioli et al. 2001	IBS	55	61	33,3	NR	28	61	33,7	10	x		
Posserud et al. 2004	IBS	25	72	43,6	NR	24	72	35,9	7		x	
Racciatti et al. 2001	CFS	36	56	38,3	88	20	56	34,2	8		x	
Riedel et al. 2002	FM	13	100	49	NR	13	100	50,1	8		x	
Riedel et al. 1998	FM	16	81	46,3	NR	17	81	39,9	8		x	
Roberts et al. 2004	CFS	56	63	39,4	56	35	63	34,9	16	x		
Roberts-Thomson et al. 1988	IBS	14	NR	47	NR	15	NR	40	3		x	x
Rowbottom et al. 1998	CFS	16	63	40,1	NR	16	63	40,5	5		x	
Scott et al. 2000	CFS	19	63	33,8	NR	10	63	27,6	8		x	
Scott et al. 1999b	CFS	15	53	40,6	NR	11	53	35,8	9		x	

Table 2 (continued). *Characteristics of the included studies in the meta-analysis.*

Study	Type of FSD	*N* of cases	% F cases	Mean age cases	Mean duration (months)	*N* of con	% F con	Mean age con	Quality (points)	Cortisol saliva	Cortisol blood	Cortisol urine
Scott et al. 1999a	CFS	13	38	38,9	60	13	38	39,4	10		x	
Scott et al. 1998c	CFS	20	65	32,9	NR	20	65	28,2	6		x	
Scott et al. 1998b	CFS	14	57	38,7	58	14	57	33,1	8		x	
Scott et al. 1998a	CFS	13	62	36,2	57	13	62	31	9		x	
Scott et al. 1998	CFS	21	67	36,1	NR	15	67	33,4	9			x
Shufflebotham et al. 2009	IBS	11	100	34,4	NR	10	100	31,2	8		x	
Strickland et al. 1998	CFS	14	100	36	NR	131	100	34	8	x		
Torpy et al. 2000	FM	13	100	44,9	NR	8	100	45,8	8		x	x
Van Denderen et al. 1992	FM	10	100	40,6	NR	10	100	40,6	7		x	
Van Rensburg et al. 2001	CFS	15	67	NR	NR	15	67	NR	5			
Videlock et al. 2009	IBS	44	57	40,4	NR	39	54	37,3	10	x		
Visser et al. 2001	CFS	59	67	38	NR	54	67	38	8		x	
Walter et al. 2006	IBS	24	85	41	NR	15	85	42	11	x		
Wingenfeld et al. 2007	FM	15	100	47,9	185	20	100	37,9	12		x	
Wood et al. 1998	CFS	10	60	34,9	37	10	60	34,2	11	x		
Yatham et al. 1995	CFS	11	73	NR	NR	11	73	NR	4		x	
Young et al. 1998	CFS	22	45	39	30	22	45	38	11	x		x
Zarkovic et al. 2003	CFS	9	67	35	23	39	67	37,8	12		x	

Abbreviations: CFS = chronic fatigue syndrome, con = controls, F = female, FM = fibromyalgia, FSD = functional somatic disorder, IBS = irritable bowel syndrome, NR = not reported.

Overall comparison cortisol and FSD

Figure 1 shows a forest plot of the SMD of baseline cortisol level in FSD subjects compared to healthy controls in each of the included studies. Meta-analysis revealed that cortisol levels were generally lower in FSD subjects compared to controls, but this association did not reach statistical significance (81 studies, SMD -0.07, 95% CI -0.17 to 0.04, $p = 0.241$). As expected, statistically significant heterogeneity in effect sizes across those studies was present (Q test $\chi^2 = 201$, $p<0.0001$, $I^2 = 60\%$).

Next, we performed separate meta-analyses to address the effect of the diurnal rhythm of the HPA axis on the association between cortisol and FSD. No statistically significant differences in cortisol between FSD subjects and controls in the morning measurements (60 studies, SMD -0.10, 95% CI -0.22 to 0.02, $p = 0.107$) or afternoon measurement (35 studies, SMD 0.01, 95% CI -0.14 to 0.17, $p = 0.868$) were found. In contrast, the summary effect size for 24-h UFC revealed statistically significant lower cortisol output in FSD subjects compared to controls (19 studies, SMD -0.42, 95% CI -0.67 to -0.18, $p = 0.001$). Statistically significant heterogeneity in the effect sizes was present for all analyses (Q test for morning measurements $\chi^2 = 129$, $p<0.001$, $I^2 = 54\%$; Q test for afternoon measurements $\chi^2 = 69$, $p<0.001$, $I^2 = 51\%$; Q test for 24-h UFC measurements $\chi^2 = 49$, p<0.001, $I^2 = 63\%$).

Moderator analyses

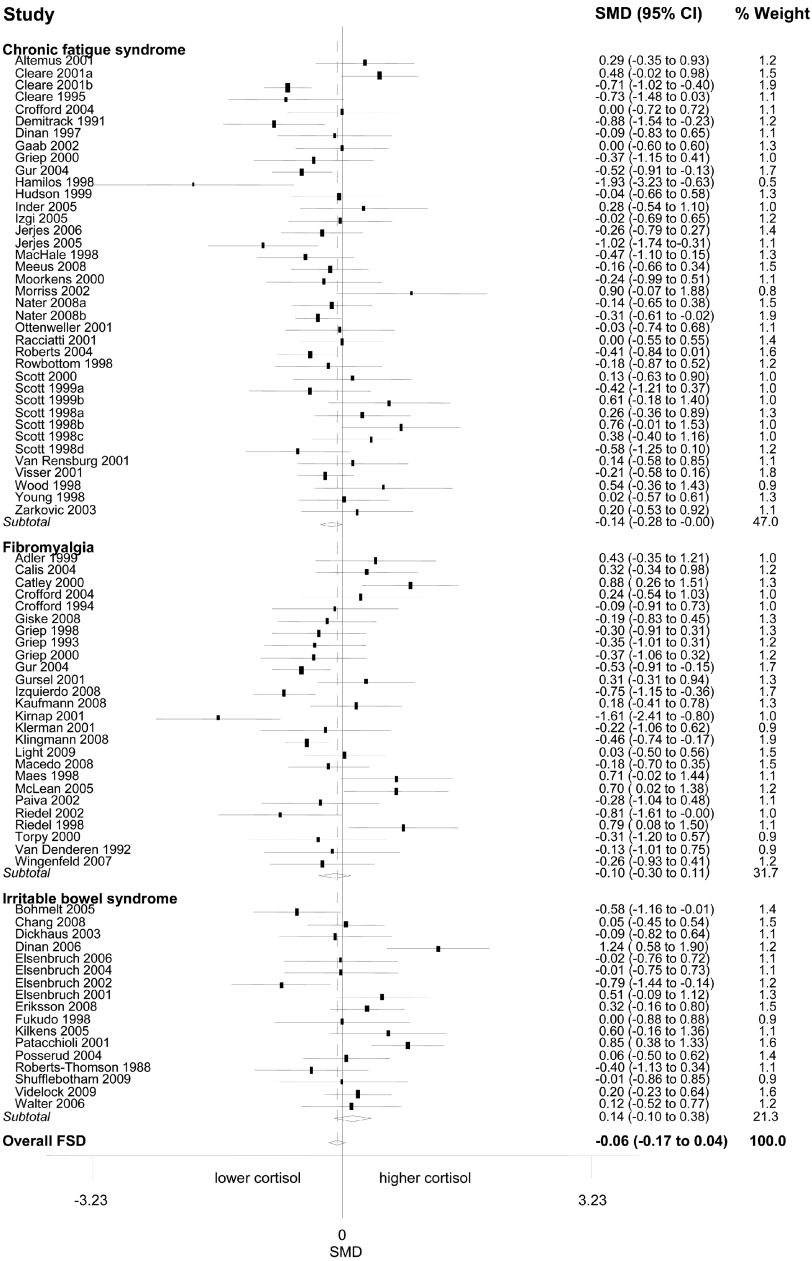
Type of FSD

We first tested whether the type of FSD influenced the summary effect size regarding cortisol levels. Statistically significant hypocortisolism was found in CFS subjects compared to controls (38 studies, SMD -0.14, 95% CI -0.28 to 0.00, $p = 0.047$). Lower cortisol levels were also found in FM subjects compared to controls; however, this difference was not statistically significant (26 studies, SMD -0.10, 95% CI -0.30 to 0.11, $p = 0.359$). Higher baseline cortisol levels were observed compared to controls in IBS patients; however, this difference was not statistically significant (17 studies, SMD 0.14, 95% CI -0.10 to 0.38, $p = 0.263$). For all three FSD, significant heterogeneity in effect sizes was present (Q test for CFS $\chi^2 = 73$, $p<0.001$, $I^2 = 49\%$; Q test for FM $\chi^2 = 71$, $p<0.001$, $I^2 = 65\%$; Q test for IBS $\chi^2 = 40$, $p<0.001$, $I^2 = 60\%$).

Gender

We divided the studies based on a median split of percentage female FSD subjects. Studies with relatively few (<77%) females did not find a difference in cortisol between FSD subjects and controls (40 studies, SMD -0.01, 95% CI -0.17 to 0.14, $p = 0.868$). Although not reaching statistical significance, studies with relatively many (≥77%) females tended to find lower cortisol in FSD subjects compared to controls (41 studies, SMD -0.11, 95% CI -0.26 to 0.04, $p = 0.132$). A post hoc exploratory analysis only including studies restricted to females, showed

Figure 1. *Forest plot of the association between cortisol and functional somatic disorders. This forest plot demonstrates standardized mean differences (SMD) and 95% confidence intervals (95% CI) of the included studies and the summary effect size of the association between baseline cortisol and functional somatic disorders (FSD).*



significant hypocortisolism in FSD female subjects compared to female controls (24 studies, SMD -0.21, 95% CI -0.37 to -0.05, p = 0.009). This effect was mainly accounted for by studies in FM (17 studies, SMD -0.24, 95% CI -0.42 to -0.06, p = 0.008) and CFS (2 studies, SMD -0.37, 95% CI -0.82 to 0.08, p = 0.109), but not IBS (5 studies, SMD -0.04, 95% CI -0.47 to 0.39, p = 0.851).

Medication use
The summary effect size approached statistical significance in studies in which medication use potentially affecting the HPA axis (corticosteroids, oral contraceptives, estrogen replacement therapy, antidepressants) was excluded (34 studies, SMD -0.16, 95% CI -0.33 to 0.06, p = 0.059), whereas there was no difference between FSD subjects and controls in studies which medication use was either not excluded or not stated as exclusion criterion (47 studies, SMD 0.08, 95% CI -0.13 to 0.15, p = 0.910).

Co-morbid depressive disorder
In studies that excluded co-morbid depressive disorder there was no difference in cortisol between FSD subjects and controls (33 studies, SMD 0.03, 95% CI -0.14 to 0.20, p = 0.730), whereas in studies that did not exclude co-morbid depressive disorder or did not state whether co-morbid depressive disorder was an exclusion criterion there was borderline statistically significant lower cortisol in FSD subjects compared to controls (48 studies, SMD -0.13, 95% CI -0.27 to 0.01, p = 0.071).

Physical inactivity
Only a minority of the studies ascertained that physical activity level in FSD subjects was comparable with controls when assessing cortisol levels (13 studies, SMD -0.05, 95% CI -0.24 to 0.14, p = 0.608). In studies that did not specifically address physical activity levels, however, the summary effect size regarding cortisol was essentially the same (68 studies, SMD -0.07, 95% CI -0.19 to 0.06, p = 0.281).

Meta-regression
Meta-regression, taking the independent effects of all above mentioned moderators into account (Table 3), demonstrates that type of FSD is a statistically significant moderator of the summary effect size, with the largest difference between CFS and IBS. A negative direction of the regression coefficient in this meta-regression indicates that presence of the moderator is associated with more hypocortisolism in FSD compared to controls. Percentage of females is also a significant moderator of the summary effect size, such that including females leads to more marked hypocortisolism in FSD subjects compared to controls. The other moderators did not have an independent statistically significant contribution; however, the directions of the regression coefficients suggest that not matching physical activity levels results in more marked hypocortisolism in

FSD subjects compared to controls; not excluding medication use results in less hypocortisolism in FSD subjects compared to controls; and excluding subjects with co-morbid depressive disorder leads to less hypocortisolism in FSD subjects compared to controls. In a model with those five moderators, explained variance of the summary effect size is 14%. Post hoc, we tested whether differences in type of assay could explain the differences in cortisol findings between the three FSD. Adding type of assay to the meta-regression did not essentially change the results.

Table 3. *Meta-regression with effect size as dependent variable and different potential moderators of the effect size as independent variables.*

	Coefficient	95% CI	*t*	*p*-value
Type of FSD[a] FM	0.25	-0.10 to 0.60	1.44	0.156
IBS	0.39	0.08 to 0.70	2.52	0.014*
Percentage females[b]	-0.01	-0.02 to 0.00	-2.00	0.049*
Medication use not excluded	0.15	-0.08 to 0.37	1.31	0.193
Exclusion of co-morbid depressive disorder	0.19	-0.04 to 0.41	1.64	0.106
Controls not matched on physical inactivity	-0.15	-0.47 to 0.17	-0.93	0.356
Adjusted r^2	0.14			

** denotes the coefficient is significant at 0.05 level.*
Abbreviations: FSD = functional somatic disorder, CFS = chronic fatigue syndrome, FM = fibromyalgia, IBS = irritable bowel syndrome.
[a] *CFS is the reference category; overall significance of FSD is* $F_{(2,72)} = 3.22$, $p = 0.046$.
[b] *Percentage females is entered as a continuous variable in this regression model.*

Sensitivity analyses

Missing values

Imputing conservative effect sizes (SMD = 0.00) for four studies that only stated that cortisol levels in FSD subjects were not significantly different from controls did not essentially change the overall summary effect size (85 studies, SMD -0.06, 95% CI -0.17 to 0.04, $p = 0.238$).

Methodological study quality

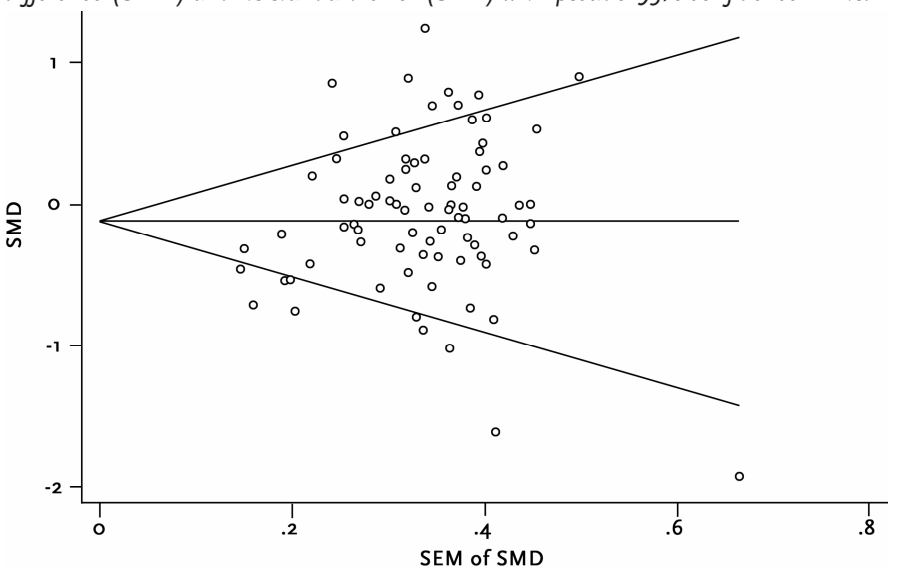
Out of the maximum of 18 points, mean quality score was 10 (range 3 - 17). The interrater reliability of methodological study quality was good ($\kappa = 0.82$). Based on a median split (low quality = less than 10 points, high quality = 10 points or more), the summary effect size for low quality studies (35 studies, SMD -0.13, 95% CI -0.30 to 0.03, $p = 0.108$) was larger but not statistically different from the summary effect size of high quality studies (46 studies, SMD -0.02, 95% CI -0.16 to 0.13, $p = 0.836$). Furthermore, heterogeneity in both quality subgroups was

comparable (*Q* test χ^2 = 74, p<0.001, I^2 = 54% and *Q* test χ^2 = 127, p<0.001, I^2 = 64%, respectively).

Publication bias

Finally, we tested whether publication bias could have affected the results. In discordance with the funnel plot from which no visual asymmetry is apparent (Figure 2), Egger's test suggested that there was significant funnel plot asymmetry (p = 0.011). However, after performing the trim and fill procedure, no studies were trimmed or filled, indicating absence of substantial publication bias.

Figure 2. *Funnel plot showing the correlation between the standardized mean difference (SMD) and its standard error (SEM) with pseudo 95% confidence limits.*



DISCUSSION

The aims of this study were to assess whether FSD are characterized by HPA axis alterations and, if present, to examine by which variables this association is moderated. A meta-analysis of 85 studies demonstrated that baseline cortisol levels were not significantly different in FSD subjects as a whole compared to healthy controls, but were significantly lower in CFS when FSD were considered separately. Female gender, medication use, and presence of co-morbid depressive disorder were moderators of this association. However, meta-regression indicated that the only independent statistically significant factors explaining heterogeneity in effect sizes of cortisol were type of FSD and female gender. It is important to note that the magnitude of the effect size in CFS and FM is comparable, with a

wider confidence interval for FM studies possibly related to a lower total number of subjects. In studies composed of exclusively female patients, hypocortisolism is also significant in FM. Thus, hypocortisolism is found in CFS and possibly FM, but not in IBS.

Several explanations can be offered to clarify why, in contrast to our hypothesis, hypocortisolism is only present in CFS and FM but not in IBS. A first explanation is that this finding demonstrates that CFS and FM are etiologically more alike than CFS or FM with IBS (Aaron et al. 2000, Sullivan et al. 2002). If CFS and FM indeed share hypocortisolism as a causal risk factor, this might also contribute to the experience of fatigue and widespread pain, which are prominent features of both conditions, but not of IBS. This explanation is, however, not fully compatible with the often advanced idea of HPA axis dysfunction as a mediator between psychosocial stress and FSD, since there are no indications that chronic psychosocial stress is differentially associated with the three FSD.

Another explanation is that cortisol levels in IBS patient may differ from those observed in FM and CFS because of the type of studies which are performed in IBS patients. Baseline cortisol measurements in IBS subjects are often performed before sigmoidoscopy or rectal extensions, both stressful procedures which may elicit acute anticipatory stress responses with a hyperactive HPA axis (Walter et al. 2006, Miller et al. 2007).

A third explanation is that cortisol alterations are not FSD-specific, but are instead specific for certain subgroups among FSD, which have the highest prevalence in CFS. A subgroup may be formed by subjects with co-morbid depressive disorder, which indeed has a higher prevalence in CFS compared to IBS (Henningsen et al. 2003). Given the association of depressive disorder with elevated cortisol levels (Burke et al. 2005, Vreeburg et al. 2009a), we expected that a meta-analysis restricted to studies excluding participants with depressive disorder would lead to more marked hypocortisolism in FSD subjects. In contrast, however, meta-analysis of studies that did not exclude co-morbid depressive disorder showed more marked hypocortisolism in FSD subjects compared to controls, whereas meta-analysis of studies that did exclude co-morbid depressive disorder did not show any differences between FSD subjects and controls. Although a potentially counterintuitive result, it could be explained by FSD patients experiencing more atypical features of depression (American Psychiatric Association 1994), given that atypical depression is characterized by hypocortisolism (Gold et al. 1995b, Antonijevic 2006). Alternatively, excluding patients with co-morbid depressive disorder may lead to the exclusion of the more severe cases of FSD, which might reduce the chance of finding HPA axis disturbances. Females could constitute another relevant subgroup across different FSD. Although only reaching statistical significance in exploratory analysis, lower cortisol levels in FSD subjects are predominantly found in studies which included a larger proportion of women.

This effect is especially apparent in FM studies (effect size in FM studies with only females is two to three times as large as in the total group of FM studies), but not in IBS studies, whereas the number of CFS studies was too low to draw conclusions. However, in a high quality, large population-based study of CFS subjects and controls included in this meta-analysis, attenuated morning salivary cortisol concentrations were only found in female CFS subjects but not in male CFS subjects (Nater et al. 2008a). Future studies are advised to consider gender as a moderator of the association between HPA axis activity and FSD, as meta-regression also indicates that a difference in percentage of included female subjects accounts for differences in cortisol levels between FSD.

Finally, FSD-specific cortisol alterations may arise from differences in behavioral consequences of the FSD, such as changes in medication use, physical activity, sleeping pattern, working status, or smoking habits (Luger et al. 1987, Badrick et al. 2007, Ambrogio et al. 2008, Vreeburg et al. 2009b). The importance of considering medication use was confirmed by our meta-analysis, although the information that could be obtained from the original studies is too general to draw firm conclusions on the exact influence of oral contraceptives, corticosteroids, estrogen replacement therapy, and antidepressants separately. Although we are not aware of studies directly assessing differences in medication use between FSD, the prevalence of antidepressant use may differ by FSD due to differences in co-morbidity with depressive disorder (Henningsen et al. 2003), or by a differential evidence base for antidepressant use for indication other than depressive disorder (e.g., pain) (Henningsen et al. 2007, Uceyler et al. 2008). We also examined the effect of physical inactivity in our meta-analysis. Given the nature of their symptoms, CFS and FM subjects might avoid physical exercise to a larger degree than IBS subjects do. Although this may theoretically explain differences between the FSD, we found no support for a role of physical inactivity in explaining differences in cortisol levels. However, it should be recognized that the robustness of this analysis is limited, as only a few studies - in particular in CFS - took this potential moderator into account. Moreover, information on physical inactivity was usually based on self-report, while using an objective method to assess physical inactivity (actigraphy), only a subgroup of the CFS patients can be labeled as persistently inactive (van der Werf et al. 2000). Although some studies have indeed shown that hypocortisolism in FSD is reversible by treatment, (Bonifazi et al. 2006, Roberts et al. 2009b) these studies have not specifically examined whether these alterations are related to reducing adverse behavioral consequences. It should be noted that all these mechanisms are probably not mutually exclusive in their contribution to differences in cortisol levels in FSD, and future studies are needed to determine which explanation carries most weight in whom.

Some limitations of this meta-analysis specifically and this research field in general should be recognized. First, this research field lacks a gold standard how

best to study the HPA axis. Several different measurement procedures are available, interpretation of which is not always unambiguous, making the research field prone to focusing on isolated false positive findings. In this meta-analysis, for example, we observed the importance of time of day of the measurement, as hypocortisolism in FSD seem especially present in the morning samples (although not reaching statistical significance in the overall-analysis), but not in afternoon samples. This relative importance of morning cortisol levels is underlined by the finding of statistically significant lower cortisol 24-h UFC in FSD subjects compared to controls, to which morning urine is important because the amount of cortisol excreted following the morning peak of HPA axis activity makes a substantial contribution to the total amount of cortisol excreted in a day (Edwards et al. 2001). These findings suggest that future studies on HPA axis activity in FSD could at least obtain the cortisol awakening response or a morning cortisol sample. Second, this study contains a systematic and quantitative analysis of robust alterations in basal cortisol levels only. Although spontaneous cortisol secretion has been considered most relevant for understanding disease processes (Nicolson 2007), systematically analyzing HPA axis dysfunction after challenge tests might help to gain a more complete picture of the HPA axis in FSD. However, the low number of available studies assessing the HPA axis after challenge tests in combination with the abundance of different tests does not allow reliable meta-analysis at this time. Another limitation is that some variables that might theoretically be important moderators could not be tested in this meta-analysis because they were seldom measured in the included studies. For example, acute psychosocial stress and psychosocial stress in the past are rarely addressed, but both impact cortisol levels (Dickerson & Kemeny 2004, Miller et al. 2007, Michaud et al. 2008). In this perspective, one relevant stressor might be physical or emotional maltreatment during youth, as it has recently been found that decreased cortisol responses to awakening are observed only in those individuals with CFS who reported exposure to childhood trauma but not in individuals without such exposure (Heim et al. 2009). Another example of an often mentioned potential moderator that could not be tested is sleep disturbance (Buckley & Schatzberg 2005). However, while subjects with FSD perceive and report sleeping disturbances significantly more often than control subjects, objective sleeping abnormalities are often absent (Elsenbruch et al. 1999, Majer et al. 2007).

We conclude that FSD cannot be collectively referred to as hypocortisolemic disorders, as this meta-analysis only confirmed the presence of lower cortisol levels in subjects with CFS and in females with FM, but not in IBS. Given the likely multifactorial etiology of FSD, HPA axis dysfunction may be of clinical relevance, especially because there are indications for the existence of subgroups in which effects sizes are substantially larger. Although it is important to realize that predictors of remission might differ from predictors of disease onset, randomized

controlled trials have shown that low-dose cortisol replacement therapy can produce short-term reductions in fatigue and other features of CFS (McKenzie et al. 1998, Cleare et al. 1999). Furthermore, a recent study shows that hypocortisolism and a flattened diurnal release of cortisol are associated with a poorer response to cognitive behavioral therapy in CFS (Roberts et al. 2009a). This implies that a patient's neuroendocrine profile may be relevant in choosing the optimal treatment strategy.

This meta-analysis provides a robust assessment of the presence of HPA axis activity alterations in FSD and the role of several potential moderators of this relationship. Several sensitivity analyses confirmed the validity of the findings. Due to the large extent of reliance on cross-sectional case-control studies in this field, however, important questions about the role of cortisol alterations in FSD remain. Cross-sectional studies are unable to shed light on the important question as to whether observed endocrine disturbances primary and causal, secondary and consequential, or epiphenomenal and causally unrelated to CFS. Nevertheless, knowledge about moderators derived from those cross-sectional studies in combination with the criteria listed in the quality tool should inform the conduct of well-designed prospective studies and aid further progress in this field.

Appendix A. *Development of quality tool for studies measuring hypothalamic-pituitary-adrenal axis activity in functional somatic disorders.*

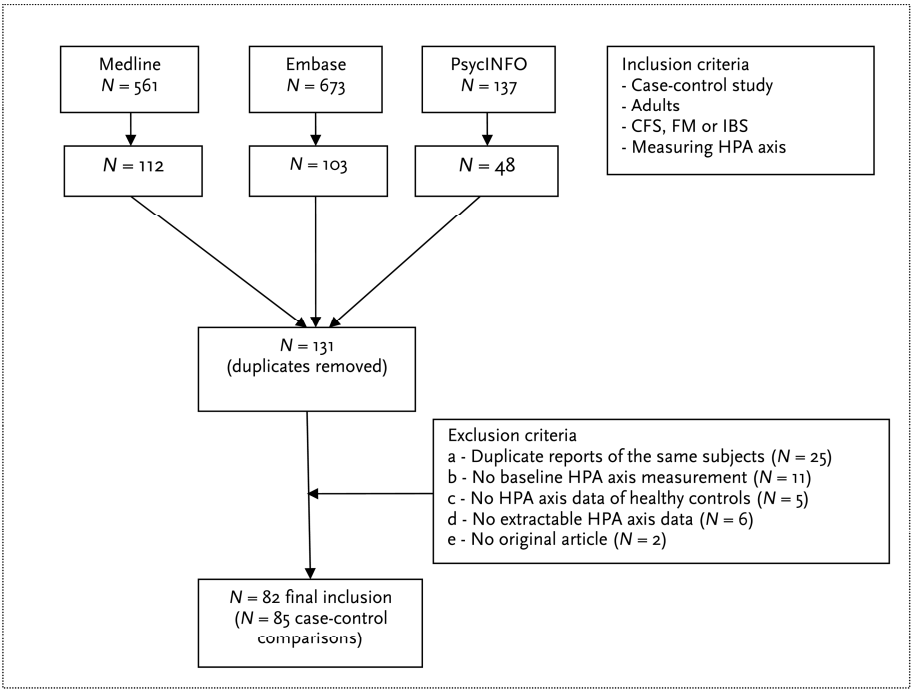
Key domain 1 (item 1-4): Appropriate selection of participants
Subjects have to meet the international consensus criteria for the respective FSD available at the time of the study (Center for Disease Control and Prevention Criteria for CFS, American College of Rheumatology Criteria for FM, Rome Criteria for IBS). Because self-report may introduce bias and underlying pathology has to be excluded, only physician diagnosis is considered valid. Controls have to represent the population from which the cases arose. Recruiting healthy controls from hospital staff or medical students may lead to selection bias (Lee et al. 2007). Somatic conditions with endocrine pathology and psychiatric conditions such as depressive disorder have been associated with altered HPA axis function (Burke et al. 2005, Nicolson 2007, Vreeburg et al. 2009a). Medication use that influences the HPA axis or interferes with peripheral assays is important to consider. Not only hormonal medication such as corticosteroids, oral contraceptives and estrogen replacement therapy, but also antidepressants should be taken into account (Holsboer & Barden 1996, Ambrogio et al. 2008). Studies should report the central tendency of disease duration with an appropriate measure of dispersion, because neuroendocrine features may differ according to phase or duration of illness (Gaab et al. 2004). There is currently no tool that has been validated to indicate severity of FSD. Therefore, symptom

frequencies of the diagnostic criteria of the respective FSD, or the Chalder Fatigue Scale in case of CFS (Chalder et al. 1993), should be reported as a proxy of disease severity enabling to assess correlations with the extent of HPA axis alterations.

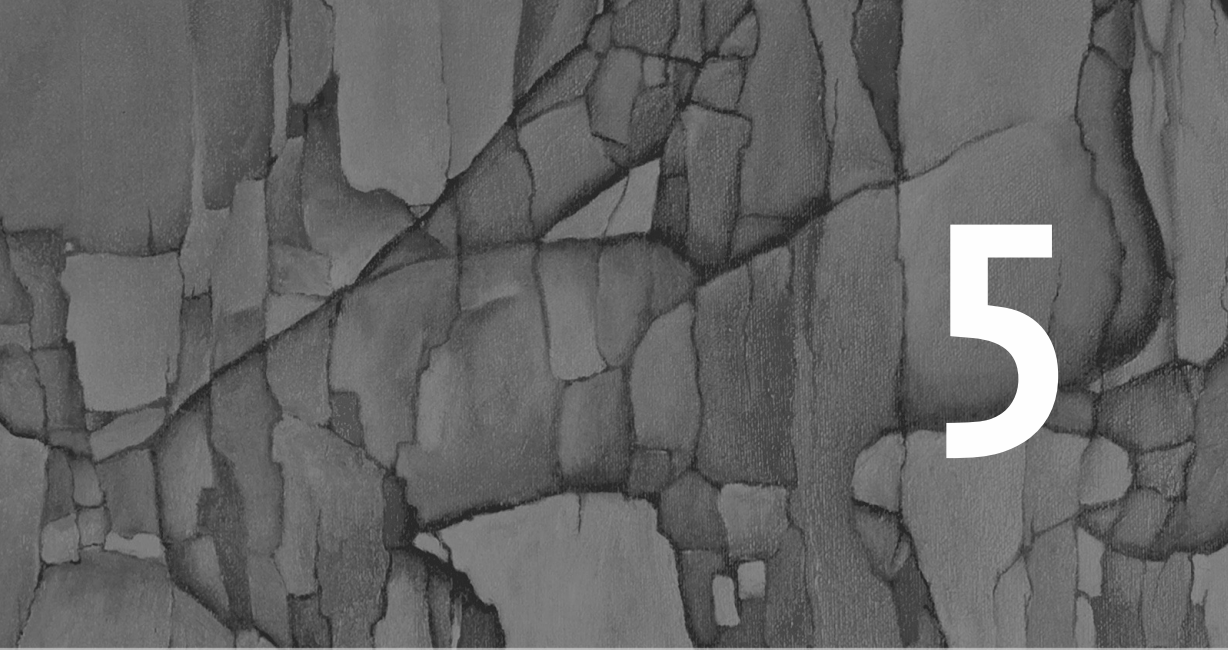
Key domain 2 (item 5-7): Appropriate quantification of HPA axis function
Blinding of researchers helps to ensure reliability, which is not only important in clinical trials, but also in experimental research (Day & Altman 2000). Researchers may approach FSD subjects differently than healthy controls during HPA axis measurements in a laboratory setting. Authors should clearly report methods of HPA axis assessment; the most important requirements are extensively reviewed elsewhere (Nicolson 2007). We selected six requirements: timing of HPA axis measurement, type of assay performed, storage conditions, calculating the mean cortisol from repeated measurements at the same time point, compliance with cortisol collection (only relevant to urinary and salivary sampling in naturalistic settings), and the behaviors shortly prior to the cortisol measurement, in particular intake of a meal or smoking, which should be restricted, standardized, or monitored (Nicolson 2007). Representation of the HPA axis outcomes should include central tendency and measures of dispersion, as well as appropriate metric or molar units.

Key domain 3 (items 8-9): Appropriate control for confounding
There are a number of potential confounders in the relation between HPA axis function and FSD. Gender (Raven & Taylor 1996), age (Van Cauter et al. 1996), smoking (Badrick et al. 2007), frequency of exercise (Luger et al. 1987), presence of depressive disorder (Burke et al. 2005), and body mass index (Ukkola et al. 2001) have all been reported to be associated with HPA axis function. Gender (Kroenke & Spitzer 1998a), age (Escobar et al. 1987), smoking (John et al. 2004), frequency of exercise (Kop et al. 2005), presence of depressive disorder (Henningsen et al. 2003), and body mass index (Neumann et al. 2008) have also been associated with FSD.

Appendix B. *Results of systematic search after applying in- and exclusion criteria and overview of excluded studies.*



Abbreviations: CFS = chronic fatigue syndrome, FM = fibromyalgia, HPA axis = hypothalamic-pituitary-adrenal axis, IBS = irritable bowel syndrome.



5

CHAPTER 5

Empirical foundations for the diagnosis of somatization: Implications for DSM-V

JGM Rosmalen, LM Tak, and P de Jonge
Submitted

ABSTRACT

The aim of this study was to develop empirically validated criteria for the diagnoses of clinically relevant somatization. This study was performed in a population cohort consisting of 461 males (47.8%) and 503 females (52.2%), with an average age of 55.7 years (SD 11.1, minimum 35.9 years, maximum 82.2 years). Somatization, anxiety, and depression were derived from the Composite International Diagnostic Interview (CIDI). Mplus was used to perform confirmative factor analyses on (1) the current DSM-IV symptom groups, (2) alternative symptom clusters previously suggested, and (3) to perform latent class analysis in order to define an empirically-derived cut-off for caseness. The existence of symptom groups as described in DSM-IV was not supported by our data, whereas a differentiation between cardiopulmonary, musculoskeletal, gastrointestinal, and general somatic symptoms did fit our data. Latent class analysis revealed two classes characterized by few (N = 884) and many (N = 80) symptoms, which could be approached by a simple cut-off of four functional symptoms (sensitivity 91%, specificity 98%, likelihood ratio 46) regardless of the number of organ systems involved. We found no empirical support for the existence of symptom clusters as currently described in DSM-IV, whereas previously suggested clustering of cardiopulmonary, musculoskeletal, gastrointestinal, and general somatic symptoms did fit our data. However, our data suggest that the diagnosis of somatization could be based on a simple cut-off of four functional symptoms, regardless of the number of symptom clusters from which they originate.

INTRODUCTION

Clusters of symptoms for which no clear organic pathology are found (i.e., functional symptoms) are classified under somatoform disorders in the Diagnostic and Statistical Manual for Mental Disorders (DSM)-IV (American Psychiatric Association 1994). The classical example of a disease characterized by numerous diverse functional symptoms is somatization disorder. However, the diagnostic criteria of somatization disorder described in the DSM classification remain largely unexplained and weakly justified. Escobar and colleagues (Escobar et al. 1998) have reviewed the criteria for somatization disorder in the subsequent DSM editions: from a lifetime history of 25 unexplained somatic symptoms in addition to attitudinal features in the original criteria; to a symptom count of 14 (males, DSM-III) or 16 (females, DSM-III) followed by 13 (both genders, DSM-IIIR); to the current criteria of 8 symptoms originating from 4 designated symptom groups in DSM-IV. Still, it has been suggested that somatization disorder is defined too exclusively, and that somatization disorder identifies only a minority of patients with clinically relevant functional symptoms.

Several elements in the diagnostic criteria for somatization disorder require further justification. First, whereas the DSM-III criteria concerned a simple symptom count, from DSM-IV onwards the diagnosis of somatization disorder required a combination of symptoms from several symptom groups (i.e., at least four pain symptoms, two gastrointestinal, one sexual, and one pseudoneurological). This criterion implies that symptoms of the different designated symptom groups cluster. Some studies, especially those performed in primary care settings or in the community, have indeed suggested that certain types of symptoms cluster (Swartz et al. 1986, Simon et al. 1996, Liu et al. 1997, Robbins et al. 1997, Gara et al. 1998, Kroenke et al. 1998b, Fink et al. 2007). However, the latent factors observed in these studies do not completely resemble the DSM-IV symptom groups: commonly identified factors are gastrointestinal, musculoskeletal, neurological/conversion, and cardiopulmonary/autonomic (Swartz et al. 1986, Simon et al. 1996, Robbins et al. 1997, Gara et al. 1998, Kroenke et al. 1998b, Fink et al. 2007), although not all studies replicated the existence of all four factors. Moreover, in several of these studies, factors were highly correlated, suggesting that a general higher order somatization factor explained the majority of the variance in functional symptoms (Liu et al. 1997, Deary 1999, Fink et al. 2007). Other studies, especially those performed in specialist care settings, did not identify any clinically meaningful symptom clusters (Hiller et al. 2001, Nimnuan et al. 2001b, Sullivan et al. 2002).

Second, it is unclear whether an appropriate symptom count threshold for caseness can be formulated. There is a continuous relationship between the number of somatic symptoms and several indicators of construct validity, including functional impairment, childhood and family risk factors, psychiatric

comorbidity, and health care use (Kroenke et al. 2007). At the same time, it may be useful to define a threshold for caseness, both from a clinical and a research point of view. In that case, it is important to have empirical justification for a cut-off.

Third, it is unclear whether the diagnostic criteria should, in addition to a symptom count threshold, also include a threshold for the number of organ systems involved. It has never been tested whether this threshold is required, nor which minimum number of organ systems would be appropriate.

Thus, there is a need for an empirically founded new approach of classifying somatization, specifically with the development of DSM–V in mind. There is agreement that physical complaints should be the focus of the classification of these disorders (Rief & Isaac 2007b), but to date it is unclear how to establish a clinical diagnosis based on these complaints. In addition, the diagnosis preferably involves recent symptoms instead of lifetime symptoms, because detailed inquiry about lifetime symptoms is typically not feasible in busy clinical settings, whereas lifetime recall of functional symptoms is also very inconsistent (Kroenke et al. 1997, Simon & Gureje 1999).

The aim of this study was to develop empirically founded criteria for the diagnoses of clinically relevant somatization, based on recent symptoms. We had the following questions. First, are there indications for symptom clusters within the functional symptoms, such as defined in the current DSM-IV classification or in some of the alternative proposals? Second, is it possible to define a cut-off for clinically relevant somatization based on population-based empirical data, and should this cut-off differ between genders? Third, is there empirical evidence to require the presence of symptoms originating from specific symptom groups, and if so, what would be an empirically supported minimum number of organ systems required? We performed this study in a population cohort using data on somatization derived from the Composite International Diagnostic Interview (CIDI).

METHODS

Population

Our study has been performed in a cohort derived from Prevention of REnal and Vascular ENd stage Disease (PREVEND), a population cohort study investigating microalbuminuria as a risk factor for renal and cardiovascular disease. The recruitment of participants has been extensively described elsewhere (Pinto-Sietsma et al. 2000). All inhabitants of the city of Groningen between the ages of 28 and 75 yrs (85421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire on demographics and cardiovascular history. A

total of 40856 subjects (47.8%) responded. After exclusion of subjects with insulin dependent diabetes mellitus and pregnant women, all subjects with an elevated urinary albumin concentration of ≥10 mg/L (N = 7768) together with a randomly selected control group with a urinary albumin concentration of <10 mg/L (N = 3395) were invited for further investigations (total N = 11163). Finally, 8592 subjects completed the total screening program, rendering the PREVEND study cohort. Because the PREVEND study population was enriched for albuminuria, this oversampling for albuminuria was counterbalanced in the current substudy. Albuminuria-negative participants and a random sample of albuminuria-positive participants were combined so that a population-representative ratio of albuminuria-positive participants was achieved. Research assistants handed over invitations to 2554 subjects to participate in the current substudy for which additional psychiatric and psychosocial data were collected during two measurement waves. Baseline measurements in the 2001 - 2002 wave were completed by a total of 1094 participants (43%). PREVEND participants who declined to participate in the current study did not significantly differ from those who did participate concerning gender (44% vs. 46%, χ^2 = 1.65, p = 0.199) and age (53.5 vs. 53.1, t = 0.88, p = 0.378). Follow-up measurements in the 2003 - 2004 wave were completed by a total of 976 participants, forming the cohort for the current study. The study was approved by the local medical ethics committee. All subjects gave written informed consent to participate in the study.

Somatization

Somatization was measured by the somatization section of the CIDI. A fully computerized version of the CIDI 2.1 12-months version was applied, suitable for self-administration. Trained interviewers were present for questions and for participants that needed computer help. In the CIDI somatization section, the occurrence of 43 symptoms in the previous year is surveyed. Symptoms are considered present when they meet severity criteria, i.e., provoke a health care visit. If these criteria are met, the interview assesses in a hierarchical fashion whether a medical doctor diagnosed a symptom as due to physical illness or injury, or whether a symptom was caused by the use of medication, drugs, or alcohol. If these inquiries are negative for medical explanations, the symptom is scored as a functional symptom. The CIDI has adequate test-retest reliability and validity (Andrews & Peters 1998). Complete CIDI data were available for 964 participants (99% of the current study cohort).

Confirmatory factor analyses

In order to test previously postulated symptom clusters, we performed confirmatory factor analyses for binary data using Mplus 3.11 (Muthen & Muthen 2004). We first tested the symptom clusters currently defined in the CIDI DSM-IV scoring algorithm: (1) pain symptoms (head (C6), abdomen (C1), back (C2), joints (C3), extremities (C4), chest (C5), genitals (C10), during menstruation (C7), during sexual intercourse (C48), during urination (C8), additional sites (C11)); (2)

gastrointestinal symptoms other than pain (nausea (C17), feeling bloated or full of gas (C19), vomiting other than during pregnancy (C15), diarrhea (C18), or intolerance of several foods (C20)); (3) sexual or reproductive symptoms other than pain (e.g., sexual indifference (C45, C49), other sexual problems including erectile or ejaculatory dysfunction (C50), irregular menses (C42), excessive menstrual bleeding (C43), vomiting throughout pregnancy (C16)); (4) pseudoneurological symptoms (e.g., conversion symptoms such as impaired balance or coordination (C24 C24A), paralysis (C26) or localized weakness (C36), difficulty swallowing or lump in throat (C41), aphonia (C27), urinary retention (C9), hallucinations (not assessed), loss of touch or pain sensation (C25A), double vision (C34), blindness (C21), deafness (C23), seizures (C28), dissociative symptoms such as amnesia (C31), or loss of consciousness other than fainting (C30)). Because of the different definition of the sexual symptom cluster in the CIDI scorings syntax rules, we performed these analyses separately for males and females. Urinary retention (C9) and difficulty swallowing or lump in throat (C41) were not reported by males, and thus not included in the analyses in males. Symptoms that are combined in the CIDI scorings algorithm (C45/C49, C24/C24A, C26/C36) were included as separate symptoms in these analyses.

In addition, we performed a confirmatory factor analysis using cardiopulmonary, musculoskeletal, and gastrointestinal factors resembling those previously reported by Fink et al. (Fink et al. 2007) and Kroenke et al. (Kroenke et al. 1998b). We also tested a four factor model, including the factor general symptoms. We defined a cardiopulmonary factor (including the items chest pains (C5), breathlessness without exertion (C35)), a musculoskeletal factor (back pain (C2), pain in joints (C3), pain in extremities (C4), loss of touch or pain sensation (C25A), localized weakness (C36), unpleasant numbness or tingling sensations (C40) and a gastrointestinal factor (pain in abdomen (C1), nausea (C17), diarrhea (C18), feeling bloated or full of gas (C19), intolerance of several foods (C20)). We also tested a four factor model, including the factor general symptoms (headaches (C6), impaired coordination (C24A), impaired balance (C24), dizziness (C29)). Finally, we performed a confirmatory factor analysis including a second order factor representing a common latent phenomenon underlying the cardiopulmonary, musculoskeletal and gastrointestinal symptom groups. The models were deemed to fit the data well if all of the following goodness-of-fit indices were satisfied: overall χ^2 goodness-of-fit test non-significant, CFI above 0.95 and RMSEA 0.05 or lower (Hu & Bentler 1999).

Latent class analyses

In order to establish an empirically derived threshold for caseness, we applied Latent Class Analysis (LCA) to the CIDI symptoms. LCA is a statistical model-fitting method identifying different classes of subjects within a given data set (Goodman 1974a, Goodman 1974b). Instead of giving a particular true solution, LCA produces solutions for different numbers of classes with relative fit indices.

The Bayesian information criterion (BIC) was used for the goodness of fit to determine the optimal number of classes. We performed separate sets of LCA for the entire cohort and for males or females separately, either including the 29 CIDI symptoms eliciting a positive response from 10 or more participants (thereby excluding pain during urination (C8), urinary retention (C9), vomiting other than during pregnancy (C15), vomiting throughout pregnancy (C16), blindness (C21), deafness (C23), impaired coordination (C24A), paralysis (C26), seizures (C28), loss of consciousness other than fainting (C30), dissociative symptoms such as amnesia (C31), bad taste in mouth, or excessively coated tongue (C38), irregular menses (C42), and pain during sexual intercourse (C48)), or including 23 CIDI symptoms after excluding the six reproductive and sexual symptoms (pain during menstruation (C7), burning sensation genitals (C10), excessive menstrual bleeding (C43), sexual indifference (C45), unpleasant sexual intercourse (C49), and other sexual problems including erectile or ejaculatory dysfunction (C50)). To index the amount to which symptoms discriminated the latent classes, we used Cramer's *V*.

Descriptive analyses and the calculation of sensitivity, specificity and predictive values of various symptom thresholds for latent class membership were analyzed using SPSS 16.0. Two-sided *p*-values below 0.05 were considered significant.

RESULTS

General characteristics

The current study cohort consists of 461 males (47.8%) and 503 females (52.2%), with an average age of 55.7 years (SD 11.1, minimum 35.9 years, maximum 82.2 years). A total of 574 participants reported at least one functional symptom, while the maximum number of reported functional symptoms was 18. A statistically significant gender difference was found in the total number of symptoms reported ($z = -6.405$, $p<0.001$). Since this gender difference might be related to the fact that the CIDI interview includes reproductive and sexual symptoms that are not equally applicable to men and women, we repeated the analysis excluding these symptoms, and found that the gender difference remained ($z = -4.884$, $p<0.001$). Besides female reproductive symptoms, significant gender differences existed for a variety of pain symptoms (back, joints, extremities, head, additional sites), dizziness, intolerance of several foods, bad taste in mouth, or excessively coated tongue, difficulty swallowing or lump in throat, sexual indifference, and unpleasant sexual intercourse, with in all cases women scoring higher. There was no association between the total number of functional symptoms and age (Spearman's rho 0.035, $p = 0.276$).

Question 1: Symptom clusters
We first performed a confirmatory factor analysis on the symptom groups that form the core of the current DSM-IV diagnostic criteria for somatization disorder. We used the CIDI scoring rules to define groups of pain symptoms, gastrointestinal symptoms other than pain, sexual or reproductive symptoms other than pain, and pseudoneurological symptoms. Because of the different definition of the sexual or reproductive symptom cluster, we performed these analyses separately for males and females. We did not find a satisfactory fit, for males (χ^2 (df 15) = 28.592, p = 0.018; CFI 0.871; RMSEA 0.044), nor for females (χ^2 (df 31) = 60.455, p = 0.001; CFI 0.856; RMSEA 0.043). Although in both cases RMSEA was below 0.05, CFI was not above 0.95 and the χ^2 test was significant.

In addition, we performed a CFA using factors resembling those previously reported (Kroenke et al. 1998b, Fink et al. 2007). Since these factors were defined based on a cohort including both males and females, we also tested them on the entire cohort. We defined a cardiopulmonary factor, a musculoskeletal factor and a gastrointestinal factor (included symptoms are described in the method section). The fit of this model was good (χ^2 (df 31) = 41.718, p = 0.095; CFI 0.965; RMSEA 0.019), and this three-factor model fitted our data significantly better than the corresponding one-factor model (χ^2 for difference testing (df 3) = 22.542, p<0.001). We also tested a four factor model, including a general symptoms factor suggested previously (Fink et al. 2007). This model also had a good fit to the data (χ^2 (df 29) = 36.658, p = 0.155; CFI 0.974; RMSEA 0.017), and again model fit was significantly better than that of the corresponding one-factor model (χ^2 for difference testing (df 5) = 20.823, p<0.001). Finally, we performed a CFA including a second order factor representing a common latent phenomenon underlying the cardiopulmonary, gastrointestinal and musculoskeletal symptom groups suggested previously (Fink et al. 2007). Also this model had a good fit to the data (χ^2 (df 31) = 41.718, p = 0.095; CFI 0.965; RMSEA 0.019).

Question 2: Empirically-based cut-off
Latent class analysis (LCA) was performed in order to identify different classes of subjects within our data set, and to test whether subjects were classified according to symptom profile or to symptom count. We performed separate LCA including either 29 or 23 symptoms (in the latter case excluding the reproductive and sexual symptoms), and for the entire cohort or males and females separately. Table 1 shows the BIC values of the LCA solutions of the different models. The best model fit (indicated by the smallest BIC value) was in all analyses achieved with a two-class model.

Table 1. *Goodness-of-fit for the Latent Class Analyses solutions.*

	BIC value					
	Symptoms with at least 10 positive responses in the total cohort*			**Symptoms with at least 10 positive responses in the total cohort, excluding six reproductive and sexual symptoms**		
Class solution	Total cohort	Females	Males	Total cohort	Females	Males
1-Class	9617.718	5830.588	3781.661	7916.660	4780.026	3166.952
2-Class	**9208.829**	**5633.950**	**3714.368**	**7524.706**	**4568.601**	**3118.669**
3-Class	9300.244	5726.175	3822.232	7591.020	4634.205	3205.513
4-Class	9457.291	5876.170	3951.135	7690.653	4723.445	3319.940

**In males excluding C7 (pain during menstruation) and C43 (excessive menstrual bleeding).*
Abbreviation: BIC = Bayesian information criterion (values represent not sample-size adjusted BIC values).

We continued with analysis on the total cohort including 23 symptoms (results for 29 symptoms are comparable and available upon request). Table 2 shows the proportion (and number) of participants in latent class 1 and 2 reporting a particular symptom. For all symptoms, the proportion positive was higher for class 1 members than for class 2 members. There were no specific symptoms characterizing class membership; participants in one of the classes displayed few symptoms ($N = 884$), and participants in the other class ($N = 80$) presented many symptoms.

We next tested whether we could approach the LCA analysis results with a simple cut-off score solely based on the total number of symptoms. Figure 1 depicts the proportion of participants in the different latent classes in relation to functional symptom count. Based on this figure, we tested sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of cut-off three, four and five for class membership, including all 43 symptoms to determine whether participants scored above or below the cut-off (Table 3). A threshold of four symptoms is the optimal cut-off in terms of sensitivity and negative predictive value: a simple cut-off score of four out of 43 symptoms correctly identified 91% of the LCA class 1 participants while correctly classifying 99% of participants with score <4 as LCA class 2. If only 23 symptoms are included and the LCA is performed in males separately, a cut-off of three symptoms performs better for males. However, if all 29 symptoms with a prevalence of more than 1% included, also in males the most optimal cut-off is four (results not shown).

Table 2. *Distribution of symptoms in latent classes.*

	Proportion (number) reporting symptom		
	Latent class 1 ***N*** **= 80**	**Latent class 2** ***N*** **= 884**	**Cramer's** ***V***
Abdominal pain	0.46 (37)	0.04 (37)	0.44
Localized weakness	0.23 (18)	0 (2)	0.43
Pain in extremities	0.49 (39)	0.06 (49)	0.41
Feeling bloated or full of gas	0.3 (24)	0.02 (16)	0.39
Difficulty swallowing or lump in throat	0.31 (25)	0.02 (21)	0.37
Loss of touch or pain sensation	0.26 (21)	0.02 (15)	0.36
Dizziness	0.38 (30)	0.04 (39)	0.35
Back pain	0.4 (32)	0.08 (67)	0.30
Joint pain	0.43 (34)	0.08 (74)	0.30
Headache	0.39 (31)	0.08 (67)	0.29
Nausea	0.13 (10)	0 (3)	0.29
Impaired balance	0.19 (15)	0.01 (12)	0.29
Shortness of breath	0.18 (14)	0.01 (11)	0.28
Intolerance of several foods	0.15 (12)	0.01 (10)	0.26
Pain during urination*	0.09 (7)	0 (2)	0.25
Numbness/tingling	0.16 (13)	0.02 (14)	0.25
Pain additional sites	0.19 (15)	0.02 (20)	0.24
Blurred vision	0.16 (13)	0.02 (15)	0.24
Chest pain	0.21 (17)	0.04 (31)	0.23
Vomiting other than during pregnancy*	0.05 (4)	0 (1)	0.19
Urinary retention*	0.03 (2)	0 (0)	0.15
Diarrhea	0.09 (7)	0.01 (12)	0.15
Impaired coordination*	0.03 (2)	0 (0)	0.15
Dissociative symptoms such as amnesia*	0.05 (4)	0 (3)	0.15
Double vision	0.06 (5)	0.01 (8)	0.13
Frequent urination	0.1 (8)	0.02 (19)	0.13
Burning sensation genitals**	0.05 (4)	0.01 (6)	0.12
Aphonia	0.09 (7)	0.02 (17)	0.12
Sexual indifference**	0.38 (30)	0.2 (176)	0.12
Other sexual problems including erectile or ejaculatory dysfunction**	0.05 (4)	0.01 (7)	0.11
Blindness*	0.03 (2)	0 (2)	0.10
Unpleasant sexual intercourse**	0.06 (5)	0.01 (13)	0.10
Irregular menses*	0.04 (3)	0.01 (6)	0.09
Seizures*	0.03 (2)	0 (3)	0.08
Excessive menstrual bleeding**	0.04 (3)	0.01 (7)	0.08
Paralysis*	0.01 (1)	0 (1)	0.07
Pain during sexual intercourse*	0.03 (2)	0 (4)	0.07
Pain during menstruation**	0.04 (3)	0.01 (10)	0.06
Skin blotches or discoloration	0.05 (4)	0.02 (16)	0.06

Table 2 (continued). *Distribution of symptoms in latent classes.*

	Proportion (number) reporting symptom		
	Latent class 1 ***N*** **= 80**	**Latent class 2** ***N*** **= 884**	**Cramer's** ***V***
Deafness*	0.01 (1)	0 (3)	0.04
Loss of consciousness other than fainting*	0 (0)	0.01 (9)	0.03
Excessively coated tongue*	0 (0)	0.01 (7)	0.03
Vomiting throughout pregnancy*	0 (0)	0.01 (6)	0.02

*Symptoms are sorted by Cramer's V. Higher values of Cramer's V indicate symptoms that better discriminated the latent classes. Including 23 CIDI symptoms (*not in model) after excluding the 6 sexual and reproductive symptoms (**excluded).*

Table 3. *Performance of different cut-off levels of functional symptoms for LCA class 1 membership.*

LCA performed in:	***N***	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**Likelihood ratio**
Total cohort						
>=3	166	1.00	0.90	0.48	1.00	10
>=4	95	0.91	0.98	0.77	0.99	46
>=5	61	0.68	0.99	0.89	0.97	68
Males						
>=3	52	0.93	0.97	0.79	0.99	31
>=4	31	0.61	0.99	0.87	0.96	61
>=5	19	0.41	1.00	0.95	0.94	NA
Females						
>=3	114	1.00	0.87	0.47	1.00	8
>=4	64	0.93	0.97	0.78	0.99	31
>=5	42	0.69	0.99	0.88	0.96	69

Question 3: The multiple organ system requirement

The DSM-IV has the requirement of symptoms coming from four designated symptom groups. We tested the distribution of symptoms over symptom clusters, using the cardiopulmonary, gastrointestinal, musculoskeletal and general symptom clusters that were shown to fit our empirical data. We found that only 10.5% of participants scoring above the cut-off reported symptoms belonging to only one symptom cluster, 46.3% reported symptoms from two, 33.3% from three and 9.5% from all four clusters. Table 4 summarizes the comorbidity with DSM-IV mood disorders. High comorbidity is found: participants scoring above the cut-off off four symptoms have a more than 5 times higher risk of having any anxiety and

depression disorder than participants scoring below the cut-off. All specific diagnoses more often occur in participants scoring above the cut-off, with one remarkable exception: the simple phobia of the blood- injection-injury type is absent in somatizers but not in controls (0.0 vs. 0.5%), whereas all other specific phobias are more prevalent in the somatizers than in the controls (animal type 2.1 vs. 0.1, natural environment type 2.1 vs. 0.8, situational type 2.1 vs. 1.0). When comparing the number of involved symptom clusters, it is evident that the enhanced comorbidity with depression and anxiety starts already at one symptom cluster, and reaches its maximum at two symptom clusters.

Figure 1. *Proportion of participants in latent classes 1 (normal line) and 2 (dotted line) in relation to functional symptom count.*

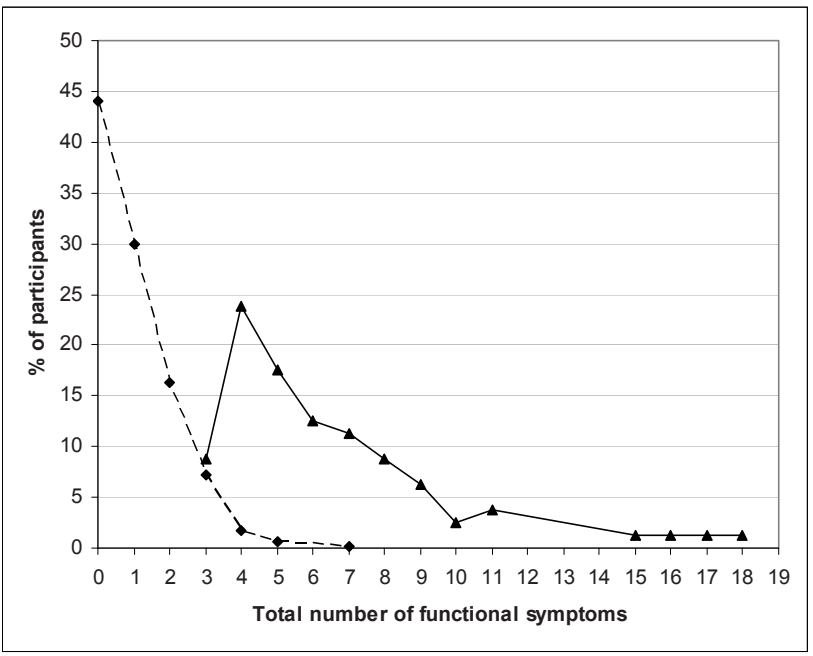


Table 4. *DSM-IV mood diagnoses in both groups %.*

	Total	<4 functional symptoms (*N* = 869)	4 or more functional symptoms (*N* = 95)				
				1 cluster (*N* = 10)	2 clusters (*N* = 44)	3 clusters (*N* = 32)	4 clusters (*N* = 9)
Any depression or anxiety disorder	10.4	7.7	38.9	20.0	40.9	40.5	44.4
Major depression	7.0	5.2	23.2	20.0	25.0	21.9	22.2
Dysthymia	0.4	0	4.2	0	2.3	6.2	11.1
Generalized anxiety disorder	2.7	1.8	10.5	20.0	9.1	9.4	11.1
Panic disorder	0.9	0.6	4.2	0	0	9.4	11.1
Agoraphobia without history of panic disorder	0.5	0.3	2.1	0	2.3	0	11.1
Agoraphobia with or without history of panic disorder	0.8	0.5	4.2	0	2.3	3.1	22.2
Social phobia	1.5	0.9	6.3	0	6.8	9.4	0
Simple phobia	2.3	2.0	5.3	0	4.5	6.2	11.1

DISCUSSION

The aim of this study was to develop empirically validated criteria for the diagnoses of clinically relevant somatization, based on recent symptoms. Our data do not fit symptom clusters as defined in the DSM-IV diagnostic criteria for somatization disorder, whereas previously suggested clusters characterized by cardiopulmonary symptoms, musculoskeletal symptoms, gastrointestinal symptoms, and general symptoms did fit our data. However, LCA revealed that a simple cut-off of 4 out of 43 functional symptoms, regardless of the number of symptom clusters from which they originate, best fits our data.

A major strength of our study is the use of a population sample including about equal numbers of both genders covering a wide age range. Moreover, we performed structured psychiatric interviews in all participants, thus without a preselection with a screening questionnaire. There are also few limitations to discuss. First, we used self-reported presence of functional symptoms, possibly underestimating the real prevalence if people tend to seek a physical reason for their complaints. In reverse, participants might have indicated that a doctor indicated a medical cause of their functional symptoms, if this cause involved a functional syndrome. Second, the utility of both factor analyses and LCA is critically dependent on the input dataset. However, despite the different interview, we were able to confirm previously suggested symptom clusters that were defined using the SCAN interview (Fink et al. 2007) and using a 15-symptom checklist derived from the PRIME-MD (Kroenke et al. 1998b).

Our data fail to provide empirical support for the designated symptom clusters in the DSM-IV. It is important to realize that we performed our analysis on symptoms experienced in the previous 12 months, whereas the DSM-IV diagnostic criteria refer to lifetime symptoms. Although earlier studies have not formally tested the DSM-IV symptom groups in confirmatory analyses, these results are in agreement with exploratory analyses that also did not find the DSM-IV clusters in a general population cohort (Liu et al. 1997). When using previously suggested symptom clusters that do fit our data, almost 90% of participants scoring above our cut-off have symptoms derived from more than one symptom cluster, suggesting that there is little additional value for a minimum required number of organ systems. Data on psychiatric comorbidity also do not give indications for such a multiple organ system threshold. Clinically significant somatization is known to be accompanied by psychiatric comorbidity (Waal de et al. 2004, Haug et al. 2004). An abrupt increase in psychiatric comorbidity beyond a certain threshold number of involved symptom clusters could thus be regarded as indicative of a dichotomy between states of health (innocent symptoms) and disease (clinically relevant somatization). Our data do not support such a symptom cluster threshold. Despite the fact that previously suggested clusters

fitted our data, LCA reveals that participants are clustered based on symptom count instead of symptom profile. This is in agreement with several earlier studies finding highly correlated symptom factors, suggesting that a general higher order somatization factor explained the majority of the variance in functional symptoms (Liu et al. 1997, Deary 1999, Fink et al. 2007).

Our data indicate that a simple cut-off of four symptoms distinguishes somatizers from non-somatizers. Although the mere counting of physical complaints as a basis for the classification has been criticized (Fink 1996), it has been shown that the number of bodily symptoms is still an important feature for the prediction of course and outcome (Jackson et al. 2006, Kroenke et al. 2007). Our results seem in agreement with results on multisomatoform disorder (MSD), which is defined as three or more currently bothersome unexplained physical complaints (from a 15-symptom checklist), plus a history of chronic somatization (i.e., unexplained symptoms that were usually present for at least two years) (Kroenke et al. 1997). Despite differences in the number of questioned symptoms and the timeframe, there are remarkable similarities between our cut-off and MSD. It is interesting that MSD disorder was present in 8% (Spitzer et al. 1994, Jackson & Kroenke 2008) to 19% (Dickinson et al. 2003) of primary care patients, compared to 9.9% scoring above our cut-off in our population cohort. In addition, psychiatric comorbidity is comparable: major depression was present in 21% of patients with MSD and 23.2% of patients above our cut-off, GAD was present in 11% with MSD and 10.5% of our somatizers, and panic disorder was present in 2% with MSD and 4.2% of patients above our cut-off (Jackson & Kroenke 2008). It appears that our LCA-based cut-off might identify the patients that fulfill the diagnostic criteria for MSD. Unfortunately, not all MSD-symptoms are surveyed in the CIDI and it is thus not possible to calculate the agreement.

Further research should validate suggested cut-offs in other samples and using other interviews. Given the fact that in some specialties functional symptoms outnumber the medically explained symptoms (Nimnuan et al. 2001a), the importance of these validations is not restricted to psychiatric settings.

6

CHAPTER 6

Age-specific associations between cardiac vagal activity and functional somatic symptoms: A population-based study

LM Tak, KAM Janssens, A Dietrich, JPJ Slaets, and JGM Rosmalen
Psychotherapy and Psychosomatics, in press

ABSTRACT

Functional somatic symptoms (FSS) are symptoms not explained by underlying organic pathology. It has been frequently suggested that dysfunction of the autonomic nervous system (ANS) contributes to the development of FSS. We hypothesized that decreased cardiac vagal activity is cross-sectionally and prospectively associated with the number of FSS in the general population. This study was performed in a population-based cohort of 774 adults (45.1% male, mean age 53.5, SD 10.7). Participants completed the somatization section of the Composite International Diagnostic Interview surveying the presence of 43 FSS. ANS function was assessed by spectral analysis of heart rate variability in the high frequency band (HRV-HF), reflecting cardiac vagal activity. Follow-up measurements of HRV-HF and FSS were performed approximately two years later. Linear regression analyses, with adjustments for gender, age, body mass index, anxiety, depression, smoking, alcohol use, and frequency of exercise, revealed an interaction of cardiac vagal activity with age: HRV-HF was negatively associated with FSS in younger adults of ≤52 yrs (β = -0.12, t = -2.37, p = 0.018), but positively with FSS in middle-aged to older adults of >52 yrs (β = 0.13, t = 2.51, p = 0.012). Longitudinal analysis demonstrated a similar pattern. Decreased cardiac vagal activity is associated with a higher number of FSS in younger adults in the general population. The unexpected association between increased cardiac vagal activity and FSS in middle-aged to older persons needs further exploration. The role of age should be acknowledged in future studies on ANS function in the etiology of FSS.

INTRODUCTION

It has been widely accepted that functional somatic symptoms (FSS), that is, physical complaints not explained by underlying organic pathology, have a multifactorial etiology with numerous contributing factors of biological, psychological, and social origin (Rief & Broadbent 2007a, Deary et al. 2007, Buffington 2009, Wise 2009). Although psychosocial stress is widely regarded as an important etiological factor (Deary et al. 2007), it is largely unknown how increased levels of psychosocial stress contribute to the experience of FSS. Dysfunction of the autonomic nervous system (ANS), a stress responsive system, is an interesting underlying mechanism to consider (Fava & Sonino 2009).

The ANS is influenced by acute, repetitive, and chronic psychosocial stress (Dishman et al. 2000, Schommer et al. 2003, Lucini et al. 2005). When the load of stressors in an individual is too large or when the ANS is chronically addressed, eventually, ANS dysfunction may develop (Chrousos & Gold 1992). Given the vital role of the ANS in the regulation of bodily organs function, increased sympathetic activity or decreased parasympathetic activity may contribute to peripheral somatosensory arousal and experience of FSS (Sharpe & Bass 1992, Rief et al. 1998, Craig 2002). Although ANS dysfunction seems a plausible mechanism to mediate the association between psychosocial stress and FSS, an alternative pathway is also possible. In this alternative pathway, ANS dysfunction may be a consequence or epiphenomenon of FSS, driven by factors such as medication use, psychiatric co-morbidity, or lifestyle factors (Tak & Rosmalen 2007). A widely used proxy for ANS function is heart rate variability (HRV), reflecting interbeat interval fluctuations in heart rate (HR). HRV indices have been particularly used to assess cardiac vagal activity as reflected in the high frequency band (HRV-HF) (Berntson et al. 1997).

ANS function has never been studied in relation to FSS in the general population. However, cardiac vagal activity has been studied in patients with functional somatic disorders (FSD), which are clusters of related FSS. Examples of FSD are chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome. A recent meta-analysis of those three FSD indicated statistically significant lower cardiac vagal activity in FSD patients compared to controls (Tak et al. 2009a), with no apparent differences between chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome. The validity of those summary estimates was, however, significantly limited by unexplained heterogeneity in the included studies. Furthermore, when taking the presence of potential publication bias into account, the difference in cardiac vagal activity between FSD patients and controls disappeared. These results suggest that some studies contrasting with the prevailing beliefs on lowered cardiac vagal activity in FSD have not been published. In addition, a structured review revealed substantial room for improvement in methodological quality, especially regarding adjustment for

potential confounders. Regarding the latter, most studies have been performed in relatively small samples and consequently lacked power to adjust for potential confounders. Furthermore, important potential moderators have not been taken into account. Only few studies have examined the associations of cardiac vagal activity and FSS according to gender (Stein et al. 2004, Tillisch et al. 2005) and none has examined the associations according to age. Especially age may be an important moderator, as recently illustrated by a study on HRV and depression. While cardiac vagal activity generally decreases with increasing age (De Meersman & Stein 2007), this study observed that cardiac vagal activity remained stable with increasing age in depressed patients but not in controls (Jindal et al. 2008). Finally, no longitudinal studies have been published exploring the issue whether ANS alterations precede development of FSS or FSD. All taken together, the current knowledge on the role of the ANS in FSS is inconclusive and requires further study.

The purpose of the present study was to investigate the role of ANS function in FSS in a large population-based cohort of adults, taking the role of potential confounders and moderators into account. We hypothesized that decreased cardiac vagal activity is cross-sectionally associated with FSS in the general population. In addition, we hypothesized that age moderated the association between FSS and cardiac vagal activity, in the sense that the association between cardiac vagal activity and FSS would be stronger in younger than in older adults. Additionally, we hypothesized that cardiac vagal activity is not differentially related to different bodily clusters of FSS. Furthermore, we hypothesized that decreased cardiac vagal activity predicts development of FSS in a 2-year follow-up period.

METHODS

Population

Our study has been performed in a cohort derived from Prevention of REnal and Vascular ENd stage Disease (PREVEND), a major population cohort study investigating microalbuminuria as a risk factor for renal and cardiovascular disease. The recruitment of participants for PREVEND has been extensively described elsewhere (Pinto-Sietsma et al. 2000). All inhabitants of the city of Groningen between the ages of 28 and 75 yrs (85421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire on demographics and cardiovascular history. A total of 40856 subjects (47.8%) responded. After exclusion of subjects with insulin dependent diabetes mellitus and pregnant women, all subjects with an elevated urinary albumin concentration of ≥10 mg/L (N = 7768) together with a randomly selected control group with a urinary albumin concentration of <10 mg/L (N = 3395) were invited for further investigations (total N = 11163). Finally, 8592 subjects completed the total

screening program, rendering the PREVEND study cohort. Because the PREVEND study population was enriched for albuminuria, this oversampling for albuminuria was rectified in the current substudy. Albuminuria-negative participants were combined with a random sample of albuminuria-positive participants until a population-representative ratio was achieved. Research assistants handed over invitations to 2554 subjects to participate in the current substudy for which additional psychiatric and psychosocial data were collected during two measurement waves. Baseline measurements in the 2001 - 2002 wave were completed by a total of 1094 participants (43%), forming our study cohort. PREVEND participants who declined to participate in the current study did not significantly differ from those who did participate concerning gender, age, and scores on a 12-item neuroticism scale (Sanderman et al. 1991). Follow-up measurements in the 2003 - 2004 wave were completed by a total of 976 participants (89%). Drop out participants ($N = 118$) were older (mean age 56.6 SD 11.9 vs. 52.7 SD 11.2, $t = 3.66$, $p<0.01$) and more often female (66% vs. 52%, $\chi^2 = 8.13$, $p<0.01$), but did not differ significantly in baseline number of FSS or in neuroticism score. The study was approved by the local medical ethics committee. All subjects gave written consent to participate in the study.

Functional somatic symptoms

FSS were measured by the somatization section of the Composite International Diagnostic Interview (CIDI). The CIDI is a fully-structured diagnostic interview developed by the World Health Organization for use in epidemiological studies on mental disorders. The CIDI somatization section provides diagnoses of somatoform disorders according to criteria in the Diagnostic Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV). A fully computerized version of the CIDI 2.1 was applied, suitable for self-administration. Trained interviewers were present for questions and for participants that needed computer help. In the CIDI somatization section, 43 symptoms are assessed through asking "have you had" this symptom, and are considered present when they meet severity criteria, i.e., provoke a health care visit. If these criteria are met, the interview assesses in a hierarchical fashion whether a medical doctor diagnosed a symptom as due to physical illness or injury, or whether a symptom was caused by the use of medication, drugs, or alcohol. If these inquiries are negative for medical explanations, the symptom is scored as a FSS. The CIDI has adequate test-retest reliability and validity (Andrews & Peters 1998). The symptom sexual indifference was excluded from analysis since it is not surveyed in the CIDI whether this symptom provoked a health care visit. Participants first completed the CIDI lifetime version measuring lifetime FSS. A total of 1088 completed CIDIs were available at baseline. Two years later, participants were re-interviewed and completed the CIDI 12-months version, in which the occurrence of the 43 symptoms in the previous year is surveyed; 964 completed CIDIs were available at follow-up. In our main analyses, we use the sum of all FSS in the CIDI 12-

months interview, henceforth defined as FSS. New-onset FSS were identified comparing the FSS reported in the CIDI 12-months interview with those reported in the lifetime interview.

Additionally, we constructed bodily clusters of FSS, based on symptom clusters previously identified in a large study on the classification of FSS (Fink et al. 2007). We defined a cardiopulmonary factor (including the items chest pain and shortness of breath), a musculoskeletal factor (including the items back pain, joint pain, pain in extremities, loss of touch or pain sensation, muscle weakness, and numbness or tingling sensations), a gastrointestinal factor (including the items abdominal pain, nausea, diarrhea, feeling bloated or flatulence, and food intolerance), and a general symptoms factor (headache, trouble with balance and walking, dizziness). A dichotomous score for every bodily system symptom cluster was calculated (0 = no FSS in specific symptom cluster, 1 = one or more FSS in specific symptom cluster).

Cardiovascular measures

Analysis of beat-to-beat variations in HR provides a non-invasive method to measure autonomic function. Both at baseline and follow-up, HRV was measured with participants lying on a bench in the supine position in a quiet laboratory room, breathing spontaneously. Because of the circadian rhythm of cardiac vagal activity (van Eekelen et al. 2004a), timing of HRV measurements was standardized in the afternoon. Research assistants who measured HRV were blind to our study hypotheses. There were no restrictions in eating, drinking, or smoking in the hours prior to the measurement. Participants were encouraged to relax and asked not to move or speak during data acquisition. A pre-rest period of 10 minutes in the supine position was applied before the HRV measurement started. A cuff was fixed around the middle phalanx of the third finger on the right hand. A Portapres device (FMS Finapres Medical Systems BV, Amsterdam, the Netherlands) continuously recorded HR during 10 minutes. Segments with a duration of approximately 300 seconds were selected for spectral analysis. In case there was no appropriate segment of 300 seconds, blocks of 60 – 300 seconds were selected (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996). A beat-to-beat variation measurement provided by the Finapres is almost interchangeable with electrocardiogram in the resting supine position (Carrasco et al. 1998). Power spectral analysis of HR signals was performed with the CARSPAN software program (Mulder 1988). The measurements were checked on stationarity and corrected for artifacts. Artifacts were eliminated and the resulting gaps were linearly interpolated. When containing more than 10% interpolated HR intervals or too many artifacts, the data were considered unstable and discarded. Participants for whom no reliable HRV measurement was available (<5% at both measurements waves) were comparable to participants with reliable HRV measurements regarding sex, age, and number of FSS. The high frequency band

(HRV-HF), defined at 0.15 - 0.40 Hz, is expressed in ms^2 and mainly reflects cardiac vagal activity (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996, Berntson et al. 1997). After natural logarithm (ln) transformation, HRV-HF values are expressed in ln (ms^2). Systolic blood pressure (BP) and diastolic BP were measured on two occasions in supine position on the right arm every minute for 10 minutes with an automatic Dinamap XL model 9300 series monitor (Johnson-Johnson Medical Inc., Tampa, FL, USA). BP (in mmHg) was calculated as the mean of the last two measurements at both occasions.

Statistical analysis

Analyses were performed using the Statistical Package for Social Sciences version 15.0 (SPSS Inc, Chicago, IL, USA). Given the skewed distribution, the number of FSS and new-onset FSS were log-transformed (after transformation, skewness 0.92, kurtosis 0.14). Extreme values (>3 SD) of HRV-HF ($N = 6$ at baseline, $N = 12$ at follow-up) were excluded. We also excluded participants using antihypertensives (beta-blockers, angiotensin converting enzyme inhibitors, and angiotensin receptor blockers assessed by self-report, $N = 140$ at baseline, $N = 165$ at follow-up), because these medications exert effects on HRV (Schroeder et al. 2003). Multivariable linear regression analyses were used to assess the independent effect of HRV-HF on the number of FSS in both the cross-sectional and longitudinal analyses, and additionally on the new-onset FSS analysis (model 1). Interaction terms were created from the centered variables to avoid problems with multicollinearity (model 2). In case of statistically significant interactions, stratified analyses are presented (model 3). Standardized β's are given. Additionally, we applied multivariable logistic regression analyses to test the effect of HRV-HF on the odds of having FSS in a specific bodily system cluster. Odds ratios (OR) and 95% confidence intervals (95% CI) are presented. All regression analyses were adjusted for gender, age, BMI, anxiety, and depressive disorder (DSM-IV diagnoses as assessed by CIDI), smoking, alcohol use, and frequency of exercise, since these may be responsible for variance in ANS function (Hayano et al. 1990, Thayer et al. 2006, Cohen & Benjamin 2006, Britton et al. 2007, Rottenberg 2007) and FSS (Kroenke & Spitzer 1998a, Henningsen et al. 2003, Glass et al. 2004, John et al. 2004, Hasin & Katz 2007, Neumann et al. 2008). Furthermore, interactions of both age and gender with HRV-HF (HRV-HF x age, HRV-HF x gender) were tested because we expected that the extent to which ANS dysfunction contributes to FSS might vary between age groups and genders (Ryan et al. 1994, Tillisch et al. 2005). Longitudinal analyses were additionally adjusted for the number of lifetime FSS at baseline. All analyses were also performed after secondary exclusion of participants having self-reported cardiovascular disease or using antidepressant medication (nonselective monoamine-reuptake inhibitors, selective serotonin reuptake inhibitors, and monoamine oxidase inhibitors)

derived from the InterAction database containing pharmacy-dispensing data (Monster et al. 2002). All *p*-values <0.05 were considered statistically significant.

RESULTS

General characteristics

Table 1 shows the general characteristics of our final study population ($N = 774$). The distribution of the total number of FSS is positively skewed; the median number of FSS was 1 (interquartile range 0 - 2). The most commonly reported FSS was headache (10.7%); other prevalent FSS were joint pain (10.6%) and back pain (9.4%).

Cross-sectional analysis of HRV-HF and FSS

Cross-sectional results from multivariable linear regression analysis are presented in Table 2. In model 1, HRV-HF is not associated with the number of FSS. After entering interaction terms HRV-HF*age and HRV-HF*gender in model 2, a significant interaction between HRV-HF and age emerges, indicating that the association between HRV-HF and the number of FSS changes with increasing age. Model 3 presents a stratified analysis based on a median split of the modifying variable age. This stratification reveals that HRV-HF is inversely and significantly associated with FSS in adults ≤52 years (β = -0.12, t = -2.37, p = 0.018), whereas HRV-HF is positively and significantly associated with FSS in adults >52 years (β = 0.13, t = 2.51, p = 0.012). Additionally, we calculated the coefficients of HRV-HF in explaining the number of FSS in five age categories, demonstrating that this interaction effect is linear (age <45 years, β = -0.19, t = -2.24, p = 0.027; 45 – 50 years, β = -0.07, t = -0.96, p = 0.339; 50 – 55 years, β = 0.00, t = 0.01, p = 0.996; 55 – 60 years, β = 0.09, t = 0.99, p = 0.324; >60 years, β = 0.15, t = 2.08, p = 0.039). In Table 1, general characteristics of the two age groups based on the median split are shown.

Cross-sectional analysis of HRV-HF and bodily clusters of FSS

Next, we examined whether the association between HRV-HF and FSS was the same in different bodily clusters. Due to the low number of participants having FSS in the cardiorespiratory bodily cluster ($N = 41$), this cluster was excluded for multivariable logistic regression. Similar to the cross-sectional analyses with the total number of FSS, for all clusters borderline statistically significant HRV-HF*age interaction effects were observed. The OR for HRV-HF x age for the three bodily symptom clusters followed the same pattern and magnitude: HRV-HF x age predicted FSS in the musculoskeletal factor (OR 1.18, 95% CI 0.98 to 1.42, z = 3.06, p = 0.080), gastrointestinal factor (OR 1.21, 95% CI 0.96 to 1.54, z = 2.58, p = 0.108), and general symptom factor (OR 1.22, 95% CI 0.99 to 1.51, z = 3.32, p = 0.069).

Table 1. *General characteristics of study population.*

	All (*N* = 774)	Age ≤52 years (*N* = 403)	Age >52 years (*N* = 371)
Age (years) (SD)	53.5 (10.7)	45.2 (4.8)	62.5 (7.7)
Gender (% female)	54.9	56.1	53.6
Body mass index (kg/m^2)	25.9 (3.8)	25.6 (3.9)	26.3 (3.6)
Smoking (%)			
No smoking	79.6	75.7	83.8
1-5 cigarette(s)/day	2.7	3.2	2.2
6-10 cigarettes/day	4.3	5.0	3.5
11-15 cigarettes/day	6.2	7.9	4.3
16-20 cigarettes/day	5.0	5.5	4.6
>20 cigarettes/day	2.2	2.7	1.6
Alcohol use (%)			
No alcohol use	20.2	17.6	22.9
1-4 unit(s)/month	16.4	15.1	17.8
2-7 units/week	33.2	39.7	26.1
1-3 unit(s)/day	27.5	25.3	29.9
>4 units/day	2.7	2.2	3.2
Frequency of exercise (%)			
Not / hardly	49.5	43.4	58.4
Once per week	29.9	28.3	21.6
Twice or more per week	20.6	28.3	20.0
Antidepressant use (%)	2.3	2.0	2.7
Depressive disorder (DSM-IV) (%)	6.7	7.7	5.7
Anxiety disorder (DSM-IV) (%)	5.7	6.4	4.7
Number of FSS, median (IQ range)			
Number of FSS	1 (0 - 2)	1 (0 - 2)	1 (0 - 2)
Number of new-onset FSS	0 (0 - 1)	0 (0 - 1)	0 (0 - 1)
Number of lifetime FSS	3 (1 - 5)	3 (1 - 5)	2 (1 - 4)
Heart rate (beats/minute), mean (SD)	69 (10)	68 (10)	69 (11)
SBP (mmHg), mean (SD)	122 (16)	118 (13)	127 (18)
DBP (mmHg), mean (SD)	72 (8)	70 (8)	73 (9)
HRV-HF			
ms^2, median (IQ range)	466 (216 - 973)	645 (322 - 1170)	314 (162 - 665)
ln (ms^2), mean (SD)	6.15 (1.28)	6.44 (1.22)	5.85 (1.28)

Abbreviations: DBP = diastolic blood pressure, DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, fourth edition, FSS = functional somatic symptoms, HRV-HF = heart rate variability in the high frequency band, IQ = interquartile, SBP = systolic blood pressure, SD = standard deviation.

Table 2. *Coefficients from multivariable linear regression models predicting the number of FSS out of HRV-HF and possible confounders: cross-sectional analysis.*

	Model 1 Number of FSS		**Model 2 Number of FSS**		**Model 3 Number FSS**			
	All participants *N* = 774		**All participants *N* = 774**		**Age ≤52 years *N* = 403**		**Age >52 years *N* = 371**	
	β	***p*-value**	***β***	***p*-value**	***β***	***p*-value**	***β***	***p*-value**
HRV-HF	0.00	0.923	-0.02	0.620	-0.11	0.018*	0.13	0.012*
Female gender	0.18	<0.001*	0.19	<0.001*	0.19	<0.001*	0.22	<0.001*
Age	0.02	0.592	0.02	0.557	0.03	0.512	0.06	0.271
BMI	0.07	0.051	0.06	0.073	0.11	0.024*	0.00	0.993
Anxiety	0.21	<0.001*	0.21	<0.001*	0.18	<0.001*	0.31	<0.001*
Depression	0.11	0.003*	0.11	0.004*	0.19	<0.001*	-0.03	0.600
Smoking	0.07	0.052	0.07	0.066	0.06	0.208	0.04	0.439
Alcohol	-0.09	0.015*	-0.08	0.024*	-0.19	<0.001*	0.06	0.259
Frequency of exercise	-0.01	0.737	-0.01	0.853	-0.02	0.691	-0.03	0.564
HRV-HF x Age			0.10	0.004*				
HRV-HF x Gender			-0.04	0.286				
r^2	0.14		0.15		0.22		0.14	

*Abbreviations: BMI = body mass index, FSS = functional somatic symptoms, HRV-HF = heart rate variability in high frequency band in ln (ms^2). *p<0.050. Model 3 is based on a median split of the moderating variable age.*

Longitudinal analysis of HRV-HF and FSS

In the longitudinal analysis, a similar pattern as found in the cross-sectional analysis emerged for the relation between HRV-HF and FSS. See Table 3 for an overview. In model 1, HRV-HF at baseline does not predict the number of FSS at follow-up approximately two years later. In model 2, the interaction term HRV-HF x age is statistically significant ($\beta = 0.07$, $t = 1.97$, $p = 0.049$). After stratification in two age groups based on a median split, the β's follow the same pattern as in the cross-sectional analyses, however, HRV-HF does not significantly predict FSS in any of those subgroups. Additionally, when only considering FSS that were new-onset, i.e., FSS that were only reported in the 12-months interview performed at follow-up but not in the lifetime interview performed at baseline, essentially the same pattern emerged (data not shown).

All analyses were also performed after secondary exclusion of participants having cardiovascular disease or using antidepressant medication. These secondary analyses did not essentially change the results; therefore, only results of the total sample are shown for maximal generalizability.

Table 3. *Coefficients from multivariable linear regression models predicting the number of FSS out of HRV-HF and possible confounders: longitudinal analysis.*

	Model 1 Number of FSS		Model 2 Number of FSS		Model 3 Number of FSS			
	All participants *N* = 700		All participants *N* = 700		Age ≤52 years *N* = 354		Age >52 years *N* = 346	
	β	*p*-value	*β*	*p*-value	*β*	*p*-value	*β*	*p*-value
HRV-HF	0.03	0.413	0.02	0.640	-0.02	0.704	0.07	0.143
Female gender	0.07	0.052	0.07	0.033*	0.10	0.049*	0.05	0.332
Age	0.09	0.006*	0.09	0.007*	0.07	0.158	0.09	0.070
BMI	0.04	0.240	0.04	0.232	0.08	0.100	0.01	0.845
Anxiety	0.09	0.014*	0.09	0.014*	0.11	0.028*	0.05	0.308
Depression	0.05	0.143	0.05	0.138	0.15	0.003*	-0.07	0.160
Smoking	0.05	0.143	0.05	0.163	0.04	0.358	0.04	0.450
Alcohol	-0.02	0.527	-0.02	0.555	-0.09	0.052	0.06	0.233
Frequency of exercise	0.00	0.965	0.00	0.967	-0.04	0.447	0.04	0.434
Lifetime FSS	0.50	<0.001*	0.50	<0.001*	0.43	<0.001*	0.56	<0.001*
HRV-HF x Age			0.07	0.049*				
HRV-HF x Gender			0.03	0.375				
r^2	0.32		0.32		0.35		0.34	

*Abbreviations: BMI = body mass index, FSS = functional somatic symptoms, HRV-HF = heart rate variability in high frequency band. *p<0.050. Model 3 is based on a median split of the moderating variable age.*

DISCUSSION

The aim of this study was to investigate the relation between ANS dysfunction and FSS in a general population cohort. As hypothesized, lower HRV-HF, which indicates decreased cardiac vagal activity, was associated with the number of FSS in adults ≤52 years. In contrast, in adults >52 years, higher HRV-HF was associated with the number of FSS. The same pattern, although not statistically significant in all analyses, emerged in the prospective part of our study when predicting FSS at 2-year follow-up.

This is the first study to report on the relationship between ANS function and FSS in a population-based cohort adjusted for a large range of confounders. In adults ≤52 years, our results are in agreement with previous findings of lower HRV-HF in rest in patients with FSD in a meta-analysis (Tak et al. 2009a). Although the reliability of the summary estimate in this meta-analysis was affected by poor quality of underlying studies, heterogeneity, and potential publication bias, the present study supports the finding of lower HRV-HF in persons with FSS. This association was the same for males and females. As hypothesized, we have demonstrated that general, musculoskeletal, and gastrointestinal FSS bodily clusters share a comparable association with HRV-HF.

Unexpectedly, as high cardiac vagal activity should protect against the sensation of FSS according to our hypothesis, we have demonstrated a significant

association between higher cardiac vagal activity and the number of FSS in adults >52 years, which remained after exclusion of participants with cardiovascular disease. Given the current lack of similar research in this age group, we can only speculate about reasons for the positive association between HRV-HF and FSS in adults >52 years. First, etiological factors in the development of FSS may differ between younger and older persons (Wijeratne et al. 2003), which may also apply to the role of stress responsive systems. In the current study, explained variance in FSS is 20% in adults ≤52 years in contrast to only 14% in adults >52 years. For example, in keeping with our results, it has been found that depressive disorder is not associated with FSS in older adults (Sheehan et al. 2004). However, age differences in etiological factors do not specifically explain the *positive* direction of the association with HRV-HF and FSS adults >52 years. Second, HRV-HF decreases with increasing age and has been associated with several diseases of aging, including obesity, diabetes mellitus, and hypertension (Masi et al. 2007), making it difficult to distinguish low HRV related to disease compared to normal aging. Also, variance in HRV-HF increases and reliability decreases with increasing age (Ergun et al. 2008). This may indicate that high HRV values reflect different physiological processes in younger persons than in older persons (Stein et al. 2005). Indeed, it has been reported that not only low HRV, as repeatedly found in previous population cohort studies, but also high HRV is a prognostic factor for cardiovascular mortality in the elderly (de Bruyne et al. 1999). Authors ascribe this remarkable finding to increased prevalence of sinus node dysfunction causing higher HRV. However, excluding participants with cardiovascular co-morbidity from our study sample did not essentially change the results. An unanticipated interaction between age and HRV-HF in psychophysiological research has been shown before, in a study examining HRV-HF and depression in the elderly. As expected, increasing age was associated with decreased HRV-HF in healthy controls; however, the depressed group did not show an association between age and HRV-HF (Jindal et al. 2008). In this study, inspection of regression lines suggests that between-group differences might be evident before age 60, but not thereafter. This cut off seems to be in line with the age-associated differences in the association between HRV-HF and FSS in the present study. Third, sensitivity to painful stimuli correlates inversely with BP levels, possible due to baroreflex mediated inhibition of pain transmission at both spinal and supraspinal levels, a process known as hypertension-associated hypalgesia (Ghione 1996). The prevalence of hypertension, which is characterized by lower cardiac vagal activity, is higher in older adults (Schroeder et al. 2003). Therefore, hypertension-associated hypalgesia would be consistent with our finding of a positive association between HRV-HF and FSS in adults >52 years. However, only a minority of the FSS concern pain symptoms. In addition, post hoc taking BP into account as a possible confounder in the current study did not significantly attenuate the positive association between HRV-HF and FSS in adults >52 years. All taken together, a comprehensive explanation for the association between

increased HRV-HF and FSS in older age groups appears is not readily available. This intriguing finding warrants further research.

The current study is the first to present prospective data on ANS function and FSS. It can be argued that alterations in cardiac vagal activity do not cause FSS, but are instead a consequence or epiphenomenon. In that case, ANS alterations are due to lifestyle factors such as physical inactivity, smoking and alcohol use, medication use, or psychiatric co-morbidity (Tak & Rosmalen 2007). However, HRV-HF was still independently associated with FSS in our analyses taking those factors into account. Longitudinal analyses, examining whether lowered cardiac vagal activity predicts development of FSS, yielded essentially the same results as the cross-sectional part of the study. Of note, prospective associations were weaker and not always statistically significant. As results in our longitudinal analyses are not conclusive, this association may be better studied in populations at risk for FSS and FSD.

Some limitations of this study should be taken into account. First, the number of FSS in the CIDI was measured by retrospective self-report. Nevertheless, we are confident that we were able to identify FSS using the CIDI, since associations with important characteristics like gender, depression, and anxiety were similar to previous studies on the prevalence of FSS in the general population (Haug et al. 2004) and in primary care (Waal de et al. 2004) in which a medical doctor was directly involved. Second, three remarks regarding our measurement of HRV should be taken into account. Although recommended as a quality criterion (Tak et al. 2009a), we did not give restrictions in eating, drinking, or smoking in the hours before the HRV measurements. However, we do not think this has influenced our results substantially, since we adjusted the analyses for smoking and alcohol use. Furthermore, we used blood pressure fluctuations as determined by the Finapres finger cuff to detect the occurrence of a heart beat. Although the distal pulse wave is almost interchangeable with an electrocardiogram (Carrasco et al. 1998), it may be argued that use of an electrocardiogram may be preferable, especially when participants with different ages are included as arterial stiffness and thereby pulse transit time is influenced by age. However, it has been shown that pulse pressure is not independently associated with HRV after controlling for age and gender. Therefore, we do not think that measuring HR peripherally has influenced our findings (Virtanen et al. 2004). Third, we did not measure respiratory rate and depth, while it has been recommended that respiration frequency needs to be monitored and adjusted for to generate an accurate measure of HRV-HF. However, this requirement seems particularly relevant when studying within subjects HRV measurements (Ritz & Dahme 2006), whereas correction or control procedures are discouraged in between subjects design such as in our study, since HRV-HF seems not dependent on respiration frequency under baseline conditions (Denver et al. 2007). Finally, we did not measure some

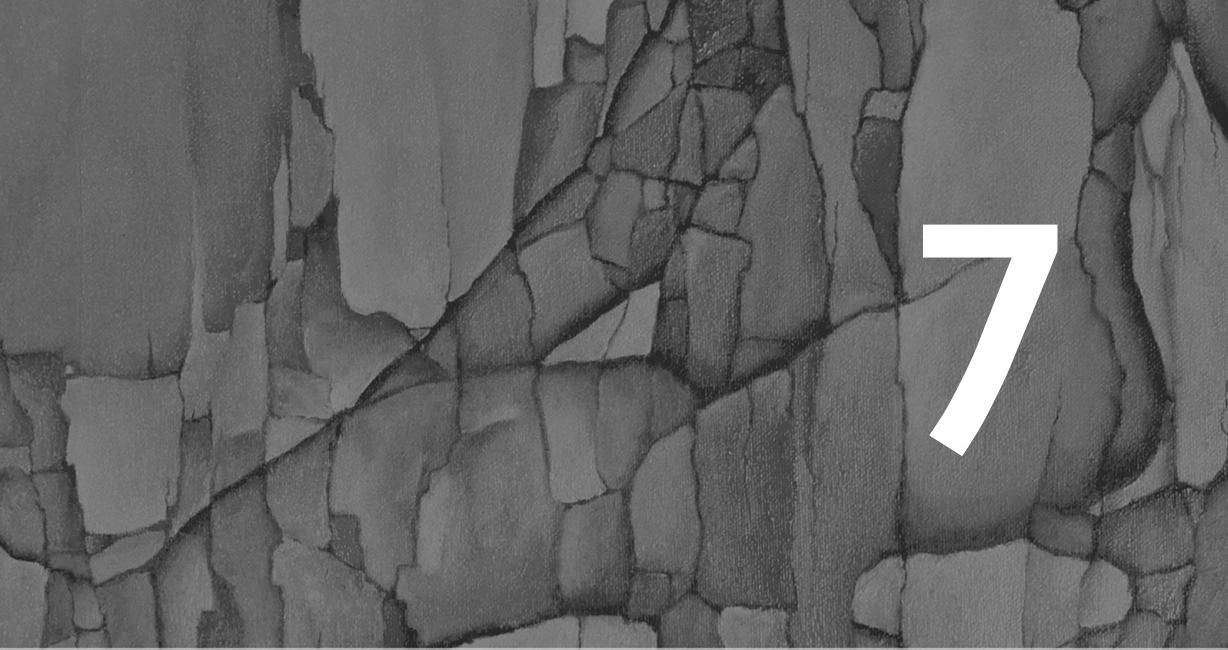
factors that may further characterize the type of somatization, such as precipitating psychological events or demoralization (Fava & Sonino 2009), and may influence its association with cardiac vagal activity.

Among the strengths of this study are the large sample size and the extensive data collection allowing for adjustment of potential confounders and detecting gender and age-dependent effects, which has been an area of weakness in previous research. In addition, our longitudinal data on ANS dysfunction in relation to FSS are unique and enabled us to study the direction of effects.

In conclusion, decreased cardiac vagal activity appeared to be associated with FSS in adults ≤52 years in the general population. Unexpectedly, higher cardiac vagal activity was associated with FSS in adults >52 years, a finding that needs replication and better understanding. Further prospective studies investigating the role of the ANS in FSS, acknowledging the role of age, are required to make a next step in clarifying the role of ANS dysfunction in the etiology of FSS.

ACKNOWLEDGMENTS

We thank ir. dr. A.M. van Roon, who provided technical assistance and expertise in heart rate variability measurements.



7

CHAPTER 7

Dysfunction of the hypothalamic-pituitary-adrenal axis and functional somatic symptoms: A longitudinal cohort study in the general population

LM Tak, SJL Bakker, and JGM Rosmalen
Psychoneuroendocrinology 2009;34(6):869-77

ABSTRACT

In persons with functional somatic symptoms (FSS), no conventionally defined organic pathology is apparent. It has been suggested that complex interactions of psychological, physiological, and social factors are involved in the etiology of FSS. One of the physiological mechanisms that may contribute to FSS is the function of the hypothalamic-pituitary-adrenal (HPA) axis. This study investigates the association of HPA axis function with cross-sectional presence and prospective development of FSS in the general population. This study was performed in a population-based cohort of 741 male and female adults (mean age 53.1 SD 10.9). Participants completed the somatization section of the Composite International Diagnostic Interview (CIDI) in which the presence of 43 FSS is surveyed. In addition to the total number of FSS, bodily system FSS clusters with musculoskeletal, gastrointestinal, cardiorespiratory, and general symptoms were constructed. HPA axis function was assessed by measuring 24-hour urinary free cortisol (24-h UFC) excretion. Follow-up measurements were performed approximately 2 years later. All analyses were adjusted for age, gender, body mass index, smoking, alcohol use, depression, exercise frequency, and urinary volume. Regression analysis detected no cross-sectional association between 24-h UFC excretion and the number of FSS (β = -0.021, t = -0.521, p = 0.603). In addition, 24-h UFC excretion was not associated with any of the bodily system FSS clusters (all p>0.050). Furthermore, 24-h UFC excretion did not predict new-onset FSS in the 2-year follow-up period (β = 0.021, t = 0.566, p = 0.572). We conclude that this study does not provide evidence for an association between altered HPA axis function, as indexed by 24-h UFC, and FSS in the general population.

INTRODUCTION

Functional somatic symptoms (FSS) are symptoms unexplained in terms of organic pathology. Functional somatic disorders (FSD) can be regarded as accumulation of persistent FSS, suggesting that they might be elicited by the same underlying mechanisms (Wessely et al. 1999). Research in this area suggests a multifactorial etiology with physiological, psychological, and social factors all contributing to FSS and FSD (Barsky & Borus 1999, Burton 2003, Deary et al. 2007).

One of the physiological mechanisms that has been suggested to contribute to FSS and FSD is the function of the hypothalamic-pituitary-adrenal (HPA) axis. Hypocortisolism is often discussed as vulnerability factor in the etiology of FSD (Heim et al. 2000a, Ehlert et al. 2001), mainly because it is regarded as a consequence of exposure to trauma and chronic psychosocial stress (Miller et al. 2007), which, in turn, may precede and contribute to development of FSD (Deary et al. 2007).

There are several hypothetical pathways how altered concentrations of corticotrophin-releasing hormone (CRH), adrenocorticotropic hormone (ACTH), and cortisol can be involved in FSD (Heim et al. 2000a). CRH not only modulates the endocrine response, but also influences pain perception. Although acute stress is known to produce analgesia, chronic stress may have the opposite effect, a process mediated by CRH (Lariviere & Melzack 2000). Low cortisol levels may cause widespread pain and fatigue (Fries et al. 2005). Alternatively, it is also postulated that HPA axis alterations may be a consequence of FSD instead of a cause, given the many factors that interact with the HPA axis, such as inactivity, psychiatric comorbidity, and medication use (Cleare 2003). When hypocortisolism indeed is secondary, it may be an epiphenomenon not linked to symptom experience at all, or a perpetuating factor in a vicious circle (i.e., exacerbating factors that had initially contributed to hypocortisolism) (Roberts et al. 2009b).

Previously, we reviewed cross-sectional studies about the HPA axis in FSD, such as chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome (Tak & Rosmalen 2007). The results of nearly 60 studies did not exclusively support one specific kind of dysfunction. The palette of different findings about the association between the HPA axis and FSD points to reduced cortisol concentration in those patients; however, results are heterogeneous with normal or high cortisol concentrations also observed. Longitudinal studies about the capacity of HPA axis activity to predict onset of FSS or FSD are interesting but scarce, not conclusive, and concern specific subpopulations (Candy et al. 2003, Rubin et al. 2005, Ter Wolbeek et al. 2007, McBeth et al. 2007).

Twenty-four hour urinary free cortisol (24-h UFC) excretion is a practical index of integrated plasma free cortisol in psychobiological studies (Levine et al. 2007). Measurement of 24-h UFC excretion has the advantage of being unaffected by the circadian rhythm of the HPA axis, is relatively easy to attain in a large cohort, and is reliable (Kushnir et al. 2003). Furthermore, collecting 24-h UFC has proven its value in previous research in detecting lower cortisol excretion in FSD patients compared to controls (McCain & Tilbe 1989, Griep et al. 1998, Cleare et al. 2001a, Cleare et al. 2001b, Jerjes et al. 2006b).

This study prospectively investigates HPA axis function and FSS in a large population cohort, allowing adjusting for a large range of possible confounders. We hypothesize that low 24-h UFC excretion is cross-sectionally associated with a higher number of FSS in the general population. Second, we hypothesize that low 24-h UFC excretion is able to predict which participants will develop new FSS in a two-year follow-up period. Additionally, we explored whether 24-h UFC excretion is differentially related to different bodily clusters of FSS.

METHODS

Population

Our study has been performed in a cohort derived from Prevention of REnal and Vascular ENd stage Disease (PREVEND), a major population cohort study investigating microalbuminuria as a risk factor for renal and cardiovascular disease. The recruitment of participants for PREVEND has been extensively described elsewhere (Pinto-Sietsma et al. 2000). All inhabitants of the city of Groningen between the ages of 28 and 75 years (85421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire on demographics and cardiovascular history. A total of 40856 subjects (47.8%) responded. After exclusion of subjects with insulin dependent diabetes mellitus and pregnant women, all subjects with an elevated urinary albumin concentration of ≥10 mg/L (N = 7768) together with a randomly selected control group with a urinary albumin concentration of <10 mg/L (N = 3395) were invited for further investigations (total N = 11163). Finally, 8592 subjects completed the total screening program, rendering the PREVEND study cohort. Because the PREVEND study population was enriched for albuminuria, this oversampling for albuminuria was rectified in the current substudy. Albuminuria-negative participants were combined with a random sample of albuminuria-positive participants until a population-representative ratio was achieved. Research assistants handed over invitations to 2554 subjects to participate in the current substudy for which additional psychiatric and psychosocial data were collected. Measurements were completed by a total of 1094 participants (43%), forming our study cohort. PREVEND participants who declined to participate in the current study did not

significantly differ from those who did participate concerning gender, age, and scores on a 12-item neuroticism scale (Sanderman et al. 1991). All subjects gave written consent to participate in the study, which was approved by the local medical ethics committee.

Functional somatic symptoms

FSS were measured by the somatization section of the Composite International Diagnostic Interview (CIDI). The CIDI is a fully structured diagnostic interview developed by the World Health Organization for use in epidemiological studies on mental disorders. The CIDI somatization section provides diagnoses of somatoform disorders according to criteria in the Diagnostic Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV). The CIDI emphasizes clinically relevant symptoms in contrast to surveying all subjective health complaints. A fully computerized version of the CIDI 2.1, suitable for self-administration, was applied. Trained interviewers were present for questions and for people that needed computer help. In the CIDI somatization section, 43 symptoms are assessed through asking "have you had" this symptom. In the CIDI, symptoms are considered present when they meet severity criteria, i.e. provoking a health care visit. If these criteria are met, the interview assesses in a hierarchical fashion whether a medical doctor diagnosed a symptom as due to physical illness or injury, or whether a symptom was caused by the use of medication, drugs, or alcohol. If these inquiries are negative for medical explanations, the symptom is scored as a FSS. The CIDI has adequate test-retest reliability and validity (Andrews & Peters 1998). Sexual indifference was excluded from analysis since it is not surveyed in the CIDI whether this symptom provoked a health care visit. Participants first completed the CIDI lifetime version. Two years later, participants were re-interviewed and completed the CIDI 12-months version, in which the occurrence of the 43 symptoms in the previous year is surveyed.

A sum score of the number of FSS was constructed summing all FSS in the previous 12 months (1-year FSS, theoretical range 0 - 42). Although we perform our main analyses on continuous FSS sum scores, we also constructed a dichotomous FSS score in order to provide information about demographic and lifestyle characteristics in groups relatively high and low on FSS. This dichotomous FSS sum score was calculated based on a latent class analysis indicating that a cut-off of participants having ≤3 FSS and participants having ≥4 FSS best fits our data on FSS. New-onset FSS were identified comparing the FSS in the CIDI 12-months version with those in the lifetime interview. Additionally, because the presence of HPA-alterations may differ between the types of FSS, we constructed three bodily system FSS clusters using factors resembling those based on a factor analysis in a large patient sample (Fink et al. 2007). We defined a cardiopulmonary factor (including the items chest pain and shortness of breath), a musculoskeletal factor (including the items back pain, joint pain, pain

in extremities, loss of touch or pain sensation, muscle weakness, and numbness or tingling sensations), and a gastrointestinal factor (including the items abdominal pain, nausea, diarrhea, feeling bloated or flatulence, and food intolerance). We also constructed a general FSS cluster based on the general symptoms factor (headache, trouble with balance and walking, and dizziness). A dichotomous score for each bodily system symptom cluster was calculated (0 = no FSS in specific symptom cluster, 1 = one or more FSS in specific symptom cluster). Finally, a dichotomous score for development of new-onset FSS for each symptom cluster was created.

Medication use

Information on drug use was obtained from the InterAction Database, containing pharmacy-dispensing data from community pharmacies in the city of Groningen. Dutch patients usually register at a single community pharmacy and therefore this pharmacy can provide an almost complete listing of the subject's prescribed drugs (Monster et al. 2002). We extracted information on drug prescriptions from 100 days prior until 100 days after the date of the visit to our research facilities. For the cross-sectional part of the study, we excluded participants using inhalation, local, gastrointestinal, or systemic corticosteroids (N = 167, 15.3%) from our analysis. For the longitudinal part of the study, participants using corticosteroids during the baseline measurement and / or during the second measurement wave were excluded (N = 275, 25.1%). Antidepressant use may influence HPA axis activity (Pariante et al. 2004, Schule 2007). Therefore, we performed all main analyses with and without exclusion of antidepressant medication users.

HPA axis function

Sampling of 24-hour urinary free cortisol (24-h UFC) occurred at two time points: at baseline and approximately 2 years later. Participants were asked to collect urine samples in a polypropylene container on two consecutive days prior to the visit to the outpatient clinic. They were instructed to urinate into the container during the 24-hour collection period and refrigerate the sample until delivery to the laboratory. UFC was measured by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Use of LC-MS/MS is recommended because this method offers advantages over immunoassays methods; the LC-MS/MS method is free of interferences from cortisol metabolites and conjugates and also eliminates drug interferences (Taylor et al. 2002). 24-h UFC was calculated by multiplying urinary volume with cortisol concentration and is expressed in nmol per 24-h. We used the mean of the two samples on consecutive days to reflect HPA axis function. Additionally, urine creatinine concentrations (mmol/L) and body surface area (m^2) (DuBois & DuBois 1916) were assessed. Repeating our analyses with 24-h UFC/urine creatinine ratio or 24-h UFC/body surface area ratio

as independent variable did not alter our results. Therefore, we used unadjusted 24-h UFC excretion in our analyses.

Potential confounders and moderators

Gender (Raven & Taylor 1996), age (Van Cauter et al. 1996), smoking (Badrick et al. 2007), alcohol use (Thayer et al. 2006), frequency of exercise (Luger et al. 1987), depression (Burke et al. 2005), body mass index (Ukkola et al. 2001), and urine volume (Mericq & Cutler 1998) may be responsible for variance in 24-h UFC excretion. Gender (Kroenke & Spitzer 1998a), age (Escobar et al. 1987), smoking and alcohol use (Hasin & Katz 2007), frequency of exercise (Glass et al. 2004), and depression (Henningsen et al. 2003) are also associated with FSS and somatoform disorders. Although the association between body mass index and FSS has never been studied, we decided to include this as a potential confounder because of its association with various mental disorders (Petry et al. 2008). Urine volume might differ between participants with and without FSS due to differences in fluid intake. Therefore, gender, age, smoking, alcohol use, frequency of exercise, depression, body mass index, and urinary volume were considered as possible confounders. Furthermore, we decided to study interactions of both age and gender with 24-h UFC (24-h UFC x age, 24-h UFC x gender) because we expected that the extent of mechanisms contributing to development and experience of FSS might differ between age groups and genders. Diagnosis of depressive disorder according to DSM-IV was derived from the depressive disorders section of the CIDI. Weight and length were measured and body mass index was calculated as the ratio between weight and the square of height (kg/m^2). Urine volume was defined as the mean of the two complete urine collections (L/24-h). The other potential confounders and additional demographic characteristics were assessed by written self-report.

Statistical analysis

Analyses were performed using the Statistical Package for Social Sciences version 15.0 (SPSS Inc, Chicago, IL, USA). Outliers of 24-h UFC were detected with the Hampel identifier (Wilcox 2002) and removed ($N = 69$, 7.7% at baseline, $N = 63$, 7.8% at follow-up). It should be noted that this outlier group was not different from the remaining study population regarding important demographic characteristics and number of FSS. The number of 1-year FSS and new-onset FSS required log-transformation to normalize their positively skewed distributions. The differences between the cut-off groups of participants having ≤ 3 FSS and participants having ≥ 4 FSS are measured with *t*-tests for parametric variables (age, body mass index, urinary volume), Mann Whitney *U* tests for nonparametric variables (sum scores of FSS of which median and interquartile ranges are presented), and chi-squares for categorical variables (gender, smoking, alcohol use, antidepressant use, presence of depressive disorder, frequency of exercise, and demographic variables). Multivariable linear regression analysis was applied to test whether 24-h UFC was cross-sectionally associated with the number of 1-

year FSS and to test whether 24-h UFC predicts occurrence of new-onset FSS in the 2-year follow-up period. Standardized β's are given. Multivariable logistic regression analysis was applied to test the effect of 24-h UFC on the odds of having FSS in a specific bodily system symptom cluster. Odds ratios (OR) and 95% confidence intervals (95% CI) are presented. We performed some secondary analyses to assess robustness of our results. All linear and logistic regression analyses were adjusted for gender, age, depression, body mass index, smoking, alcohol use, frequency of exercise, and urinary volume. Interaction terms between 24-h UFC and age (24-h UFC x age) and 24-h UFC and gender (24-h UFC x gender) were created from the centered original variables to avoid problems with multicollinearity. All *p*-values less than 0.050 were considered statistically significant.

RESULTS

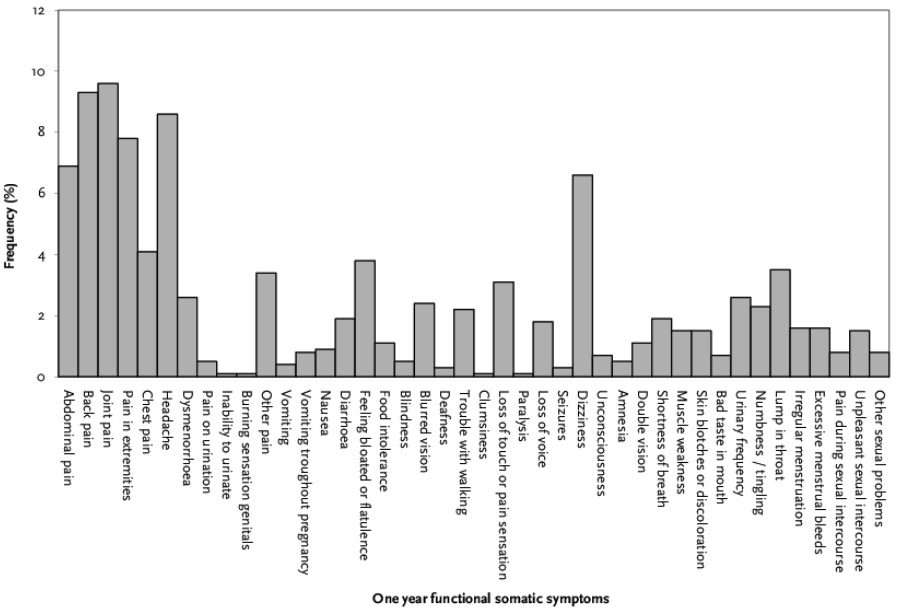
General characteristics

In the overall population, the number of reported FSS in the CIDI was positively skewed. The median number of 1-year FSS was 0 (interquartile range 0 - 1 FSS). Females reported more FSS than males (females median 1 FSS (interquartile range 0 - 2 FSS), males median 0 FSS (interquartile range 0 - 1 FSS), $z = -4.768$, $p<0.001$). The most common 1-year FSS was joint pain (9.6%), closely followed by back pain (9.3%), and headache (8.6%). Figure 1 depicts the frequency of all 1-year FSS in the total study population. The frequency of female-reproductive symptoms reflects the frequency in females only. For our main analyses, we measured the number of 1-year FSS on a continuous scale. However, for illustrative purposes, we provide general characteristics and demographic information of the study population as a whole and for participants having ≤3 FSS and participants having ≥4 FSS separately (see Table 1).

Cross-sectional analysis of cortisol and 1-year FSS

Linear regression analysis revealed that there was no association between 24-h UFC and the number of 1-year FSS ($\beta = -0.021$, $t = -0.521$, $p = 0.603$). Furthermore, logistic regression did not reveal differences in respective 24-h UFC excretions of extreme groups, that is participants with ≥4 1-year FSS ($N = 49$) and participants without any 1-year FSS ($N = 405$) (OR 0.970, 95% CI 0.532 to 1.769, $z = 0.010$, $p = 0.970$). Next, we tested whether 24-h UFC was differentially related to one of the bodily system symptom clusters. No association was found between the number of 1-year FSS in any of the other bodily system symptom clusters on the one hand and 24-h UFC on the other hand (all $p>0.050$). See Table 2 for an overview.

Figure 1. *Frequency of functional somatic symptoms (FSS) in 1-year version of Composite International Diagnostic Interview (CIDI) in the total study population.*



Categorizing FSS in bodily symptom clusters is arbitrary, as numerous slightly different classifications have been introduced. Therefore, we conducted post hoc analyses with two other symptom classifications. Underlining our main analyses, no association became apparent between 24-h UFC and (a) FSS categories of somatization disorder according to DSM-IV or (b) abridged somatization disorder FSS clusters based on the CIDI (Escobar et al. 1998). It should be noted that in the latter abridged somatization disorder bodily clusters (i.e., neurological FSS, gastrointestinal FSS, musculoskeletal FSS, and pain-related FSS) there was an association between 24-h UFC and musculoskeletal FSS (OR 1.383, 95% CI 1.012 to 1.890, $z = 4.152$, $p = 0.042$) that did, however, not remain statistically significant after Bonferroni correction for multiple testing. We also tested for potential interactions (24-h UFC x age, 24-h UFC x gender) and the possibility of a curvilinear association between 24-h UFC and the number of 1-year FSS (Loucks et al. 2008), and again, no association became apparent. Excluding participants using antidepressants from the analysis did not appreciably alter the results.

Table 1. *General characteristics.*

	All participants (*N* = 741)	Participants ≥4 FSS (*N* = 49)	Participants ≤3 FSS (*N* = 692)	Test statistic[a]	*p* value[a]
Age (years), mean (SD)	53.1 (10.9)	54.8 (11.0)	53.0 (10.9)	t = -1.091	0.275
Gender (% male)	47.6	30.6	48.8	χ^2 = 6.098	0.014*
Body mass index (kg/m²), mean (SD)	26.4 (3.8)	27.4 (3.9)	26.3 (3.8)	t = -1.891	0.059
Smoking (%)				χ^2 = 5.737	0.333
No smoking	75.6	63.3	76.5		
1-5 cigarette(s)/day	4.6	6.1	4.5		
6-10 cigarettes/day	3.5	4.1	3.5		
11-15 cigarettes/day	6.5	8.2	6.4		
16-20 cigarettes/day	5.8	10.2	5.5		
> 20 cigarettes/day	3.9	8.2	3.6		
Alcohol (%)				χ^2 = 8.065	0.089
No alcohol use	19.0	30.6	18.2		
1-4 unit(s)/month	17.9	24.5	17.4		
2-7 units/week	34.0	26.5	34.6		
1-3 unit(s)/day	24.7	16.3	25.1		
> 4 units/day	4.5	2.0	4.7		
Exercise frequency (%)				χ^2 = 1.071	0.585
Not / hardly	50.7	57.1	50.2		
Once per week	28.0	26.5	28.1		
Twice or more per week	21.3	16.3	21.7		
Antidepressant use (%)				χ^2 = 9.970	0.002*
NSMRI	0.3	2.0	0.2		
SSRI	1.5	8.1	0.9		
All	2.0	8.2	1.6		
Depressive disorder (DSM-IV) (%)	5.9	26.5	4.5	χ^2 = 39.837	<0.001*
Household composition (%)				χ^2 = 5.276	0.260
Living alone	28.0	32.7	27.7		
Living with partner and children	30.3	22.4	30.7		
Living with partner	35.8	32.7	36.1		
Living with children	3.6	8.2	3.3		
Other	2.3	4.1	2.2		
Education (%)				χ^2 – 16.406	0.001*
Low	26.3	30.4	26.0		
Middle	26.7	37.0	26.0		
High	42.5	19.6	44.1		
Not applicable	4.5	13.0	3.9		
24-h UFC (nmol), mean (SD)	66.9 (35.5)	65.3 (32.6)	70.0 (35.7)		
Urine volume (L), mean (SD)	1.8 (0.6)	1.8 (0.6)	1.8 (0.6)	t = -0.221	0.825
Number of FSS, median (IQ range)					
Number of 1-year FSS	0 (0 - 1)	5 (4 - 6)	0 (0 - 1)	z = -12.904	<0.001*
Number of new-onset FSS	0 (0 - 1)	2 (1 - 4)	0 (0 - 1)	z = -11.149	<0.001*
Number of lifetime FSS	3 (1 - 5)	7 (5 - 11)	2 (1 - 4)	z = -8.079	<0.001*

Abbreviations: DSM-IV = Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, FSS = functional somatic symptoms, IQ = interquartile range, NSMRI = nonselective monoamine-reuptake inhibitor, SD = standard deviation, SSRI = selective serotonin reuptake inhibitor.

*[a] Participants having ≥4 FSS versus participants having ≤3 FSS. * p<0.05.*

Table 2. *Multivariable logistic regression analysis on the association of 24-h UFC and three bodily system clusters of 1-year FSS.*

Predictor	Outcome bodily clusters		
OR (95% CI)	Musculoskeletal FSS	Gastrointestinal FSS	General FSS
24-h UFC	1.143 (0.851 – 1.534)	0.772 (0.514 – 1.161)	1.021 (0.726 – 1.438)
Female gender	2.147 (1.461 – 3.156)*	1.098 (0.671 – 1.799)	1.783 (1.137 – 2.797)*
Age	1.002 (0.985 – 1.019)	1.003 (0.981 – 1.025)	0.990 (0.970 – 1.010)
BMI	1.026 (0.797 – 1.074)	1.049 (0.984 – 1.113)	0.951 (0.899 – 1.007)
Smoking	1.125 (0.987 – 1.284)	0.986 (0.824 – 1.179)	1.007 (0.861 – 1.178)
Alcohol use	0.980 (0.832 – 1.153)	0.954 (0.770 – 1.183)	0.859 (0.711 – 1.039)
Depression	2.644 (1.381 – 5.066)*	3.092 (1.463 – 6.532)*	2.118 (1.022 – 4.389)*
Exercise frequency	0.874 (0.693 – 1.102)	0.867 (0.637 – 1.181)	0.665 (0.500 – 0.885)*
Urinary volume	1.152 (0.826 – 1.606)	1.030 (0.656 – 1.617)	1.457 (0.998 – 2.128)

*List of abbreviations: BMI = body mass index, FSS = functional somatic symptoms, OR = odds ratio, 95% CI = 95% confidence interval. * p<0.05. OR for 24-h UFC are adjusted for gender, age, body mass index, smoking, alcohol use, depression, frequency of exercise, and urinary volume (OR of these potential confounders are also shown). The bodily cluster of cardiorespiratory FSS has not been shown because the low prevalence of having FSS in this cluster did not allow logistic regression.*

Longitudinal analysis of cortisol and new-onset FSS

Linear regression analysis did not reveal an association between baseline 24-h UFC excretion and development of new-onset FSS (β = 0.021, t = 0.566, p = 0.572). Applying logistic regression to extreme groups, that is participants with ≥2 new-onset FSS (N = 92) and participants without any new-onset FSS (N = 503), did not reveal differences in respective 24-h UFC excretions (OR 0.961, 95% CI 0.750 to 1.230, z = 0.100, p = 0.752). Finally, only the number of new-onset FSS in the musculoskeletal cluster was large enough to allow for logistic regression analysis; having new-onset FSS in this group (12% of the total study population) was not predicted by 24-h UFC at baseline (OR 1.112, 95% CI 0.891 to 1.387, z = 0.877, p = 0.349). Excluding participants using antidepressants did not alter these findings.

DISCUSSION

The aim of the current study was to investigate the association between HPA axis function and the number of FSS, adjusting for the effects of age, gender, depression, body mass index, smoking, alcohol use, frequency of exercise, and urinary volume. Firstly, we did not find a cross-sectional association between the excretion of 24-h UFC and the total number of reported 1-year FSS nor with any of the bodily system clusters of FSS. Furthermore, 24-h UFC excretion did not predict development of new-onset FSS in a 2-year follow-up period.

Previous studies on HPA axis function and FSS have usually included clinical samples with a specific FSD, like chronic fatigue syndrome, fibromyalgia, or irritable bowel syndrome. Some of those studies did not find an association between 24-h UFC and FSD (Young et al. 1998, Adler et al. 1999, Crofford et al. 2004). However, the predominant cross-sectional finding of lower cortisol concentrations in urine, saliva, or serum in FSD patients (Cleare 2003, Fries et al. 2005, Tak & Rosmalen 2007) contrasts with findings in the present study. Several explanations may account for these conflicting results. First, no gold standard of HPA axis function testing in FSS and FSD exists. As a consequence, measurements of salivary, serum and urinary cortisol in rest as well as after challenge tests have been studied in order to draw conclusions about HPA axis function. One can question comparability of these various methods for assessing HPA axis function; achieving consensus about which HPA axis alterations might be clinically relevant in FSS and FSD is vital. Second, in the initial phase of having FSS or a FSD, other HPA axis alterations may be present than in the chronic course. The importance of chronicity is illustrated by a study in chronic fatigue syndrome patients, which observes that the extent of the HPA axis dysregulation correlates with the duration of symptoms (Gaab et al. 2004). Furthermore, it is unclear whether observed HPA axis changes in previous studies are primary and possibly causally related, or that they are secondary to risk factors for chronicity of FSD, such as physical inactivity and psychiatric co-morbidity (Cleare 2003). Heterogeneity in previous study samples concerning presence and adjustment for those factors may explain differences in cortisol findings. Third, analyzing a population cohort may also explain why we did not find an association as demonstrated in some previous studies investigating more severely affected clinical populations. This explanation is supported by two cross-sectional studies which detected larger HPA axis alterations in the most severely affected patients. One study that examined HPA axis function in fibromyalgia patients and patients with unexplained low back pain reported that HPA axis alterations in low back patients were less pronounced than those in patients with fibromyalgia (Griep et al. 1998). Another study demonstrated that those at risk of chronic widespread pain have less marked HPA axis abnormalities than the group with established chronic widespread pain (McBeth et al. 2005). A study that recruited participants from websites on several FSD, thus probably positioned between our population-

based cohort and clinical samples regarding severity of the FSS, detected no differences in salivary cortisol values between patients with FSS and healthy controls (Houtveen & van Doornen 2007b). Thus, HPA axis dysfunction may be only associated with FSS in the more severe diagnostic range.

The main outcome measures in our study are the total number of all 1-year or new-onset FSS. It has been frequently thought that different symptoms might be caused by different stress responsive system dysfunctions. For instance, gastrointestinal FSS may be accounted for by hyperactivity of the autonomic nervous system whereas musculoskeletal FSS may be attributed by HPA axis dysfunction (Fink et al. 2007). In the current study, however, 24-h UFC was not differentially associated with FSS in specific bodily clusters.

No longitudinal studies on the association between HPA axis function and FSS in the general population exist. Moreover, unlike cross-sectional studies, longitudinal studies on the association between HPA axis function and FSD are scarce. The majority of longitudinal studies available concern functional somatic fatigue that supports our null-findings (Candy et al. 2003, Rubin et al. 2005, Ter Wolbeek et al. 2007). However, in contrast to the studies on functional somatic fatigue, a longitudinal study on functional widespread pain did find an association between HPA axis function and FSS. Among a group of subjects at high risk for chronic widespread pain based on their psychosocial profile, HPA axis function appeared to predict new onset of chronic widespread pain, a condition related to fibromyalgia (McBeth et al. 2007). This study requires replication, as it may indicate a different role for the HPA axis in musculoskeletal FSS compared to other FSS.

One should consider some limitations when interpreting our results. First, we measured the number of FSS by the CIDI, a retrospective interview based on self-report without independent verification by medical records. Data about lifetime symptoms should be interpreted with caution because of measurement errors due to inaccurate recall (Leiknes et al. 2006). Even considering 1-year data, it remains unknown whether physicians have explained the nature of the symptoms accurately to the patients and whether patients have agreed and recalled the physicians' opinion correctly. For example, headache diagnosed as tension-type headache or abdominal pain diagnosed as irritable bowel syndrome may be considered as physical illness by the participant. In that case, the CIDI will score the symptom as explained by a physical illness and not as FSS. We expect that this limitation will result in an underestimation of the true amount of FSS in our population, since patients presumably tend to interpret FSS as medically explained rather than the opposite (Robins et al. 1982). However, we are confident that we have been able to measure FSS with our approach, as characteristics like gender ratio and prevalence of depression are similar to studies using a physician to decide whether symptoms are unexplained (Speckens et al. 1996, Waal de et al.

2004). Precise estimates of prevalence of FSS are strongly dependent on the population under study and definition of FSS. Due to the CIDI requirement of a health care visit and the lack of a conventional medical condition explaining the symptom, prevalence of FSS in our study is substantially lower than in population studies that survey all subjective health complaints (Ihlebaek et al. 2002). Second, we used 24-h UFC excretion to assess HPA axis function. Disadvantages include 24-h UFC being an end product of the HPA axis, through which information about the complete HPA axis, like CRH, ACTH, diurnal rhythm abnormalities, or expression of corticosteroid receptors is not available.

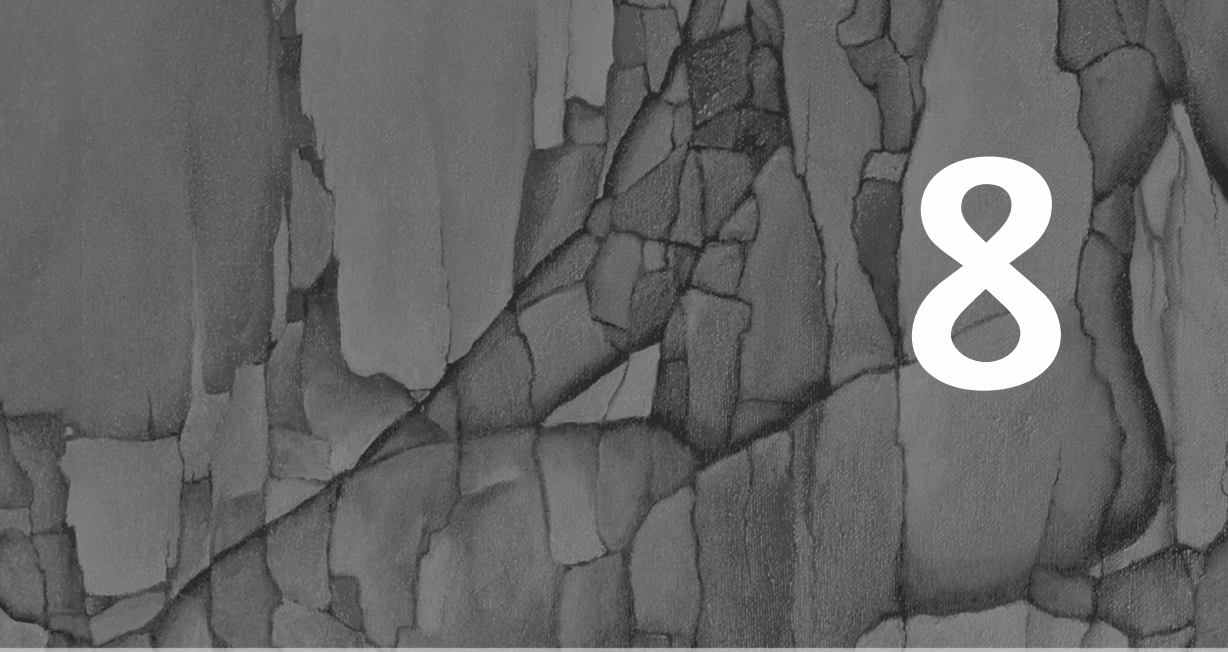
Strengths of our study include operationalising FSS in a dimensional approach using a continuous measure instead of exclusively comparing cases and controls. This dimensional operationalisation of FSS better fits daily practice, in which no widely accepted, clear-cut diagnosis exists (Kroenke et al. 2007). Furthermore, generalizability of our results is good, because we used a large population cohort without applying strict inclusion criteria.

The etiology of FSS is frequently understood as a mere cognitive-attributional process, overlooking the importance of psychobiological factors. Although we did not find a relation between HPA axis function as measured by 24-h UFC and FSS, alternative psychobiological mechanisms that may contribute to developing or experiencing FSS should be considered (Rief & Barsky 2005). Stress responsive systems, such as immune system dysfunction (Dimsdale & Dantzer 2007) and autonomic nervous system dysfunction (Tak & Rosmalen 2007), or more central mechanisms such as sustained arousal that may facilitate the development of sensitization to bodily signals (Ursin 1997, Eriksen & Ursin 2004, Brosschot et al. 2006) are worth further unraveling.

In summary, low 24-h UFC excretion is not cross-sectionally associated with a higher number of 1-year FSS in the general population. Furthermore, low 24-UFC excretion does not predict development of new-onset FSS in a 2-year follow-up period. We conclude that this study does not provide evidence for our hypothesized association between altered HPA-function and FSS in the general population.

ACKNOWLEDGEMENTS

We thank Sipke Visser, who provided expertise in and data on medication use and Jacko Duker for technical expertise regarding 24-h UFC measurements.



8

CHAPTER 8

Is high-sensitive C-reactive protein a biomarker for functional somatic symptoms? A population-based study

LM Tak, SJL Bakker, JPJ Slaets, and JGM Rosmalen
Brain, Behavior, and Immunity 2009;23(7):1014-9

ABSTRACT

Functional somatic symptoms (FSS) are symptoms unexplained in terms of underlying organic pathology. Alterations in the immune system function may be associated with FSS via induction of sickness behavior. We aimed to investigate whether low-grade immune system activation is positively associated with FSS in a population-based cohort of 881 adults (46% male, mean age 53.0, SD11.4). Participants completed the somatization section of the Composite International Diagnostic Interview surveying the presence of 43 FSS. Innate immune function was assessed by measuring high-sensitive C-reactive protein (hs-CRP). Follow-up measurements of hs-CRP and FSS were performed approximately 2 years later. Regression analyses, with adjustments for gender, age, body mass index, anxiety, depression, smoking, alcohol use, and frequency of exercise, did not reveal a cross-sectional association (β = 0.01, t = 0.40, p = 0.693) or longitudinal association (β = -0.03, t = -0.93, p = 0.352) between hs-CRP and the total number of FSS. When examining different bodily clusters of FSS, hs-CRP was not associated with the gastrointestinal FSS cluster, but the association approached statistical significance for the general FSS cluster (OR 1.08, 95% CI 0.98 to 1.18) and musculoskeletal FSS cluster (OR 1.08, 95% CI 0.99 to 1.17). For the latter association, exploratory analyses revealed that mainly the pure musculoskeletal complaints were responsible (OR 1.12, 95% CI 1.03 to 1.21). We conclude that the level of hs-CRP is not a biomarker for the total number of FSS in the general population. The association between hs-CRP and musculoskeletal and general FSS needs further study.

INTRODUCTION

Functional somatic symptoms (FSS), symptoms unexplained by underlying organic pathology, are multifactorially caused (Mayou & Farmer 2002). Functional somatic disorders (FSD), such as chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome, can be regarded as clusters of several FSS and are likely to share a common etiology (Wessely et al. 1999). Virtually all explanatory models on the etiology of FSS and FSD assign a role for psychosocial stress (Deary et al. 2007). One of the psychobiological mechanisms that may link psychosocial stress to FSS is the immune system (Segerstrom & Miller 2004).

Different pathways may be etiologically involved in this link. Firstly, the immune system may be related to FSS via sickness behavior, a constellation of non-specific symptoms induced by pro-inflammatory cytokines, such as interleukin (IL)-1, IL-6, and tumor necrosis factor (TNF)-α (Dantzer 2001, Wieseler-Frank et al. 2005). These non-specific sickness behavior symptoms, including fatigue, weakness, malaise, hyperalgesia, and increased focus on own body, are core characteristics of persons having FSS and FSD (Dantzer 2005). Secondly, immune activation may be a marker for activation of other stress responsive systems that are possibly involved in FSS and FSD, namely, the autonomic nervous system and the hypothalamic-pituitary-adrenal axis (Heim et al. 2000a, Tak & Rosmalen 2007), as both systems closely interact with the immune system (McEwen et al. 1997, Araujo et al. 2006).

Indeed, numerous abnormalities in immune cell quantity and function have been identified in several FSD (Klimas & Koneru 2007, Liebregts et al. 2007). However, those abnormalities are rarely replicated, which is illustrated by a critical review on immunological abnormalities in chronic fatigue syndrome (Lyall et al. 2003). Furthermore, a study combining patients with different FSD did not find differences in pro- and anti-inflammatory cytokines between patients and healthy controls (Houtveen et al. 2007a). These inconsistent findings may be due to the fact that relevant alterations in central cytokine levels, if present, are not adequately reflected in peripheral blood measurements (Dantzer 2001).

IL-6 and TNF-α have a central role in stimulating the liver to produce C-reactive protein (CRP), which might be a more integrated and accurate peripheral marker for innate immune system activation. CRP is a particularly suitable measure for large epidemiological studies because assessment is relatively easy and reference values are known. Although historically long been considered as clinically irrelevant, minor elevations of CRP (3 - 10 mg/L) have been reported to be associated with psychosocial stress (McDade et al. 2006) and a variety of somatic as well as psychiatric conditions (Kushner et al. 2006). Ultrasensitive assays can detect CRP in this subclinical range as high-sensitive CRP (hs-CRP). Of associations of circulating concentrations of CRP and pro-inflammatory cytokines

in subjects of the general population without active infection, those between IL-6 and CRP have been most frequently investigated and are generally strongest, with reported significant correlation coefficients ranging from 0.24 to 0.50 (Ridker et al. 2000, Cesari et al. 2003, Piche et al. 2005, Berrahmoune et al. 2007, Stewart et al. 2008, Milaneschi et al. 2009). In the study with the strongest correlation between IL-6 and CRP, which included 991 subjects aged 65 years and older, correlations of CRP with other investigated cytokines were lower: significant correlations of 0.09 and 0.19 were found for IL-1β and IL-18 respectively, and an non-significant correlation of 0.06 was found for TNF-α (Milaneschi et al. 2009). In an elderly population of 2225 subjects, TNF-α and CRP were found to have a weak but significant correlation (correlation coefficient 0.13) (Cesari et al. 2003). In smaller studies of 112 postmenopausal women and of 315 adults, correlations of TNF-α with CRP were non-significant (Piche et al. 2005, Berrahmoune et al. 2007).

In this study, our hypotheses are that hs-CRP is positively associated with the number of FSS in a large population cohort, and that hs-CRP elevations predict development of FSS in a two-year follow-up period. Additionally, we explore whether hs-CRP levels are differentially related to different bodily clusters of FSS.

METHODS

Study population

Our study has been performed in a population cohort in Groningen, The Netherlands. The selection procedure and sample characteristics have been extensively outlined elsewhere (Tak et al. 2009b). Baseline measurements in the 2001 - 2002 wave were completed by 1094 participants (aged 33 - 79 years). Follow-up measurements in the 2003 - 2004 wave were completed by 976 participants. The study was approved by the local medical ethics committee and all subjects gave written consent to participate.

C-reactive protein

Fasting blood samples were collected in all participants during a visit to the research facilities. In case of flu or a febrile temperature, blood collection was postponed to a later time (available blood samples at baseline N = 1024, at follow-up N = 983). High-sensitive CRP (hs-CRP) was determined by nephelometry with a threshold of 0.18 mg/L and intra- and inter-assay coefficients of less than 4.4 and 5.7%, respectively (BNIIN, Dade Behring, Marburg, Germany). CRP levels below the detection level were scored as 0.18 mg/L. One-time measurement is a reliable measure of CRP over an extended period of time in healthy individuals (Macy et al. 1997). Since concentrations of CRP increase dramatically as part of acute inflammatory response, it has been recommended to

discard plasma CRP concentrations above 10 mg/L in studies on CRP as a risk factor for cardiovascular disease (Pearson et al. 2003). Likewise, CRP levels above 10 mg/L may obscure the association between CRP elevations and FSS. Therefore, we eliminated subjects with CRP levels above 10 mg/L from our analyses (N = 44 at baseline, N = 25 at follow-up).

Functional somatic symptoms

FSS were measured by the somatization section of the Composite International Diagnostic Interview (CIDI), for which the procedure has been outlined elsewhere (Tak et al. 2009b). In brief, 43 symptoms are assessed through asking "have you had" this symptom. In the CIDI, symptoms are considered present when they meet severity criteria, i.e., provoked a health care visit. If these criteria are met, the interview assesses in a hierarchical fashion whether a physician diagnosed a symptom as due to physical illness or injury, or whether a symptom was caused by the use of medication, drugs, or alcohol. If these inquiries are negative for medical explanations, the symptom is scored as a FSS. The CIDI has adequate test-retest reliability and validity (Andrews & Peters 1998). Participants first completed the CIDI lifetime version, measuring lifetime FSS (N = 1088 completed CIDIs at baseline). Two years later, participants completed the CIDI 12-months version (N = 964 completed CIDIs at follow-up), surveying the same 43 symptoms in the previous year. As a measure of FSS, we summed all FSS reported in the 12-months interview. Furthermore, we created a measure of new-onset FSS by summing the FSS that were only reported in the 12-months interview but not in the lifetime interview.

Additionally, we constructed bodily clusters of FSS, based on symptom clusters previously identified in a large study on the classification of FSS (Fink et al. 2007). We generated a cardiopulmonary cluster (including questions about chest pain and shortness of breath), a musculoskeletal cluster (including questions about back pain, joint pain, pain in extremities, loss of touch or pain sensation, muscle weakness, and numbness or tingling sensations), a gastrointestinal cluster (including questions about abdominal pain, nausea, diarrhea, feeling bloated or flatulence, and food intolerance), and a general symptom cluster (including questions about headache, trouble with balance and walking, and dizziness). For all clusters, a dichotomous score was generated, with a score of 0 if no FSS in the symptom cluster was scored present and a score of 1 if one or more FSS in the symptom cluster was scored present.

Statistical analysis

Given the skewed distribution, the number of FSS and new-onset FSS were log-transformed. We used forced entry linear regression analyses to test whether hs-CRP was cross-sectionally (using hs-CRP and FSS measured at follow-up) or longitudinally (using hs-CRP measured at baseline and FSS measured at follow-up, adjusted for the number of lifetime FSS) associated with the total number of

FSS. Additionally, we tested whether hs-CRP predicted new-onset FSS. Standardized β's are given. Logistic regression analyses were performed to test whether hs-CRP is differentially associated with FSS in bodily clusters. Odds ratios (OR) and 95% confidence intervals (95% CI) are presented. Gender, age, body mass index (BMI), depression, anxiety, smoking, alcohol use, and frequency of exercise explain variance in hs-CRP (Visser et al. 1999, Hutchinson et al. 2000, Albert et al. 2003, Bazzano et al. 2003, Liukkonen et al. 2006, Lakoski et al. 2006, Pitsavos et al. 2006, Kuo et al. 2007) and FSS (Kroenke & Spitzer 1998a, Henningsen et al. 2003, Glass et al. 2004, Hasin & Katz 2007, Petry et al. 2008, Neumann et al. 2008). Therefore, all regression analyses were adjusted for gender, age, BMI, presence of depressive disorder (previous-year diagnosis of DSM-IV depressive disorder as assessed by the CIDI), presence of anxiety disorder (previous-year diagnosis of DSM-IV panic disorder with or without agoraphobia, agoraphobia, general anxiety disorder, and social phobia as assessed by the CIDI), smoking (none, 1-5 cigarette(s)/day, 6-10 cigarettes/day, 11-15 cigarettes/day, 16-20 cigarettes/day, >20 cigarettes/day), alcohol use (none, 1-4 unit(s)/month, 2-7 units/week, 1-3 unit(s)/day, >3 units/day), and frequency of exercise (not/hardly, once per week, twice or more per week). We also present results of regression models that are only adjusted for gender and age to illustrate the influence of the covariates. Interaction terms of gender with hs-CRP (hs-CRP x gender) and age with hs-CRP (hs-CRP x age) were created. All models were evaluated for absence of multicollinearity. All analyses were repeated after exclusion of participants with medication use (corticosteroids, antihypertensives, analgesics, oral contraceptives) or somatic disease (cardiovascular disease, rheumatoid arthritis, chronic obstructive pulmonary disease, inflammatory bowel disorders, malignancy) that may influence hs-CRP levels. All *p*-values less than 0.05 were considered statistically significant.

RESULTS

At baseline, the mean age of the study population was 53.0 years (SD 11.4), with 46% males. The median hs-CRP concentration was 1.18 mg/L (interquartile range 0.55 - 2.58 mg/L). The median number of FSS in the previous year was 1 (interquartile range 0 - 2 FSS, range in population 0 – 19 FSS). Prevalence of having FSS was 6.7% (N = 60) in the cardiorespiratory cluster, 23.7% (N = 227) in the musculoskeletal cluster, 12.7% (N = 114) in the gastrointestinal cluster, and 17.4% (N = 156) in the general cluster. Of the total population, 38.3% developed at least one new-onset FSS.

In the regression models only adjusted for age and gender, no cross-sectional association between hs-CRP and the total number of FSS was found (β = 0.05, t = 1.425, p = 0.155), nor did hs-CRP predict the total number of FSS (β = -0.01, t = -

0.38, p = 0.703). After adjustment for several confounders, as shown in Table 1, hs-CRP was not associated with the total number of FSS in the cross-sectional analysis or in the longitudinal analysis. There were no statistical significant interactions of hs-CRP with gender or age (data not shown). Furthermore, hs-CRP did not predict the number of new-onset FSS (β -0.03, t = -0.76, p = 0.449).

Table 1. *Multivariable linear regression explaining the number of FSS out of hs-CRP and possible confounders.*

	Outcome number of FSS Cross-sectional analysis N = 881			**Outcome number of FSS Longitudinal analysis N = 863**		
	β	***t***	***p* -value**	***β***	***t***	***p*-value**
hs-CRP	0.01	0.40	0.693	-0.03	-0.93	0.352
Female gender	0.16	4.77	<0.001*	0.09	2.93	0.003*
Age	0.05	1.44	0.150	0.07	2.34	0.020*
Body mass index	0.03	5.74	0.431	0.04	1.19	0.235
Anxiety	0.19	5.74	<0.001*	0.12	3.90	<0.001*
Depression	0.13	3.74	<0.001*	0.07	2.11	<0.035*
Smoking	0.02	0.49	0.628	0.05	1.48	0.141
Alcohol use	-0.10	-2.95	0.003*	-0.03	-1.08	0.283
Frequency of exercise	-0.06	-1.84	0.066	-0.04	-1.44	0.151
Lifetime number of FSS				0.44	14.48	<0.001*
Adjusted r^2	0.11			0.28		

Abbreviations: FSS, functional somatic symptoms; hs-CRP, high-sensitive C-reactive protein.
**p<0.050.*

Next, we tested whether hs-CRP was differentially associated with FSS in any of the bodily clusters (see Table 2). Due to the low number of participants having FSS in the cardiorespiratory cluster, this cluster was not suitable for logistic regression analyses. In the logistic models only adjusted for gender and age, hs-CRP was significantly associated with FSS in the musculoskeletal factor. When adjusting for several confounders, hs-CRP was not associated with FSS in the gastrointestinal cluster, whereas hs-CRP tended to be associated with FSS in the general cluster (p = 0.124) and musculoskeletal cluster (p = 0.085). In an exploratory analysis, the latter association appeared to be mainly driven by the pure musculoskeletal FSS, namely, back pain, joint pain, pain in extremities, and muscle weakness (model only adjusted for gender and age OR 1.14, 95% CI 1.06 to 1.23, z = 11.53, p = 0.001, fully adjusted model OR 1.12, 95% CI 1.03 to 1.21, z = 6.80, p = 0.009) and not by the more neurological FSS in this factor. Results of the longitudinal analyses showed the same positive, although weaker association for musculoskeletal FSS, whereas the association with general FSS disappeared (Table 2).

Table 2. *Multivariable logistic regression analyses on the association of hs-CRP and bodily clusters of FSS.*

Predictor OR (95% CI)	Outcome bodily clusters Musculoskeletal FSS	Gastrointestinal FSS	General FSS
Cross-sectional analysis			
hs-CRP[a]	1.11 (1.03 - 1.20)**	1.04 (0.94 - 1.15)	1.07 (0.98 - 1.16)
Adjusted hs-CRP[b]	1.08 (0.99 - 1.17)	0.99 (0.89 - 1.10)	1.08 (0.98 - 1.18)
Longitudinal analysis			
hs-CRP[a,c]	1.07 (0.98 - 1.16)	0.99 (0.89 - 1.10)	0.97 (0.88 - 1.07)
Adjusted hs-CRP[b,c]	1.04 (0.95 - 1.13)	0.94 (0.83 - 1.06)	0.99 (0.89 - 1.10)

*Abbreviations: FSS – functional somatic symptoms, hs-CRP – high-sensitive C-reactive protein, OR = odds ratio, 95% CI = 95% confidence interval. [a] Adjusted for gender and age. [b] Adjusted for gender, age, BMI, smoking, alcohol use, depression, anxiety, and exercise frequency. [c] Adjusted for the lifetime number of FSS at baseline. **p<0.01.*

All analyses were repeated after exclusion of participants with medication use (corticosteroids, antihypertensives, analgesics, oral contraceptives) or somatic disease (cardiovascular disease, rheumatoid arthritis, chronic obstructive pulmonary disease, inflammatory bowel disorders, malignancy) that may influence hs-CRP levels. These secondary analyses did not essentially change the results; therefore, only results of the total sample are shown for maximal generalizability.

DISCUSSION

In this population-based study, we did not find evidence for an association between hs-CRP and the total number of FSS. When differentially exploring bodily clusters of FSS, hs-CRP tended to be associated with general and (pure) musculoskeletal FSS.

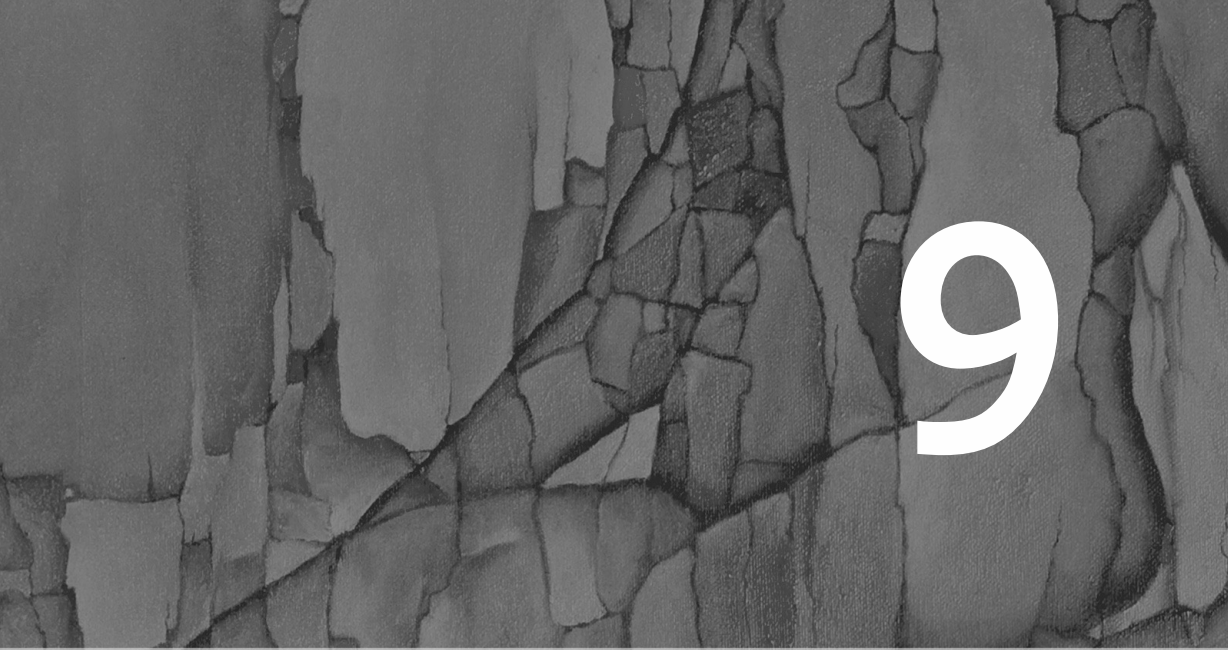
To our knowledge, no previous studies addressed hs-CRP and a large range of FSS in the general population, however, the few clinical and population-based studies that examined hs-CRP in patients with FSD may serve as a reference. In line with our findings on FSS in the gastrointestinal bodily cluster, hs-CRP levels of patients with irritable bowel syndrome were not different from those of healthy controls (Schoepfer et al. 2008). We are unaware of studies specifically focusing on musculoskeletal FSS, such as in fibromyalgia, however, hs-CRP was related to low back pain in a study in young female adults (Shiri et al. 2008) and self-reported pain not due to chronic disease in a study in older adults (Graham et al. 2006). Also in medically explained chronic pain conditions, sickness behavior is often considered to be an important contributor to symptom experience. The

association between hs-CRP and pure musculoskeletal FSS, which remained after exclusion of participants with somatic conditions such as rheumatoid arthritis, is in line with previous literature documenting an association between low-grade inflammation and chronic pain. The differential associations may be explained by the fact that those FSS are related to sickness behavior, mainly to the hyperalgesia and weakness components. This idea is supported by studies on chronic fatigue syndrome which consistently show that hs-CRP levels are higher in patients compared to healthy controls (Buchwald et al. 1997, Spence et al. 2008, Raison et al. 2009). Authors of the latter study suggest that this increase in peripheral inflammatory signaling might not be specific to chronic fatigue syndrome, but apply to unwellness in general. Alternatively, as the association of hs-CRP with musculoskeletal FSS became weaker and the association with general FSS disappeared in the longitudinal analysis, the association between hs-CRP and those FSS may be confounded by conditions or behavior that influence CRP levels, such as obesity and depression. Of note, our analyses were adjusted for several of such confounders, and hs-CRP remained independently associated with pure musculoskeletal FSS. If hs-CRP is a biomarker of those intermediate mechanisms, however, this analysis might be over adjusted. We have chosen our confounders a priori based on theory, as this strategy will generally constitute stronger scientific evidence than models that were achieved in other ways (Babyak 2004). The drawback of this strategy is inferring that hs-CRP is not causally important because its association with FSS is eliminated by the inclusion of covariates in the adjustment process, while it may only reflect that the covariates treated as confounders are actually critical to the causal chain (Christenfeld et al. 2004). For example, IL-6 and other pro-inflammatory cytokines are not only released by the liver, but have also been shown to be released by adipose tissue (Mohamed-Ali et al. 1997, Pou et al. 2007). In this perspective, BMI may not be a confounder in the potential relation between hs-CRP and FSS, but rather a causal factor inducing low-grade inflammation and sickness behavior. The same kind of conceptual problems apply to anxiety, depression, and health behaviors, such as regularly being engaged in physical activity. To further disentangle this problem, future studies could employ methods that appropriately account for the time-varying nature of the association by repeatedly assessing hs-CRP, FSS, and covariates like BMI, depression, anxiety, and health behaviors, preferably analyzed using multi-level statistical techniques (Christenfeld et al. 2004). This type of studies has more power to establish whether the association between hs-CRP and musculoskeletal FSS is causal, consequential, or epiphenomenonal.

The findings of this our study should be interpreted in light of some limitations. First, we only measured hs-CRP, which is a general biomarker for immunological activation that may not be sufficient to capture all aspects of immune function. In the introduction, we mentioned sickness behavior and activation of stress responsive systems as potential links between the immune system and FSS

(McEwen et al. 1997, Heim et al. 2000a, Dantzer 2005, Araujo et al. 2006, Tak & Rosmalen 2007). Regarding the former, it may be questioned whether hs-CRP is an adequate biomarker for sickness behavior in the general population. Current evidence for the existence of sickness behavior mainly originates from experimental research in animals and research in patients with medically explained conditions (Hart 1988, Dantzer et al. 2008a), and has focused on cytokines. Pro-inflammatory cytokines that are considered most important in sickness behavior are IL-1β and TNF-α (Dantzer et al. 2008b), while CRP is more strongly associated with IL-6 than with IL-1β and TNF-α (Ridker et al. 2000, Cesari et al. 2003, Piche et al. 2005, Berrahmoune et al. 2007, Stewart et al. 2008, Milaneschi et al. 2009). Thus, it cannot be inferred from the current results that sickness behavior does not play a role in the experience of FSS in the general population. Instead, our negative results might be explained by the fact that hs-CRP is not a valid indicator for cytokines associated with sickness behavior. Second, although the CIDI is a widely-used, validated instrument to diagnose somatoform disorders (Andrews & Peters 1998), we have no data on whether patients interpreted and recalled the physicians' opinion correctly. For example, headache diagnosed as tension-type headache or muscle pain diagnosed as fibromyalgia may be considered as physical illness by the participant. In that case, the CIDI will score the symptom as explained by a physical illness and not as FSS. We expect that this limitation will result in an underestimation of the true amount of FSS in our population, since patients presumably tend to interpret FSS as medically explained rather than the opposite (Robins et al. 1982). However, we are confident that we have been able to measure FSS with our approach, as the female preponderance and the about equally strong associations with anxiety and depression are similar to previous studies in the general population (Haug et al. 2004) and to a study in primary care having involved a physician to decide whether symptoms are unexplained (Waal de et al. 2004). Third, we did not take into account the severity of the reported FSS. If hs-CRP elevations are only associated with FSS that have reached a certain level of severity, our measurement of the total number of clinically relevant FSS could have diluted this association. The major strength of this study is its large cohort and extensive data collection, providing the possibility to adjust for several confounders.

In conclusion, we did not find evidence for a relationship between hs-CRP levels and the total number of FSS in the general population. Intriguingly, hs-CRP seems differentially related to pure musculoskeletal FSS. This finding needs further attention in future studies that are designed to gain more insight in which variables should be considered as confounders, which as mediators, and which as moderators.



9

CHAPTER 9

Effects of cumulative stress on stress responsive systems: Results from a population-based study

SL Van Ockenburg, LM Tak, SJL Bakker, ROB Gans, P de Jonge, and JGM Rosmalen
Submitted

ABSTRACT

The theory of allostatic load attributes the detrimental effects of psychosocial stress on health to a dysregulation of stress responsive systems. However, evidence that stress exposure alters the stress responsive systems function in the long run is limited. Our objective was to assess whether self-reported adverse life events during the lifespan are associated with current activity of the autonomic nervous system (ANS), the hypothalamic-pituitary-adrenal axis (HPA axis), and the immune system. This study was performed in a population-based cohort of 1094 adults aged 28 - 79 years, 46.3% male, average age 53.1. Cumulative lifetime exposure to adverse life events was assessed through the list of threatening experiences. ANS function was assessed by spectral analysis of heart rate variability in the high frequency band (HRV-HF), reflecting parasympathetic activity. HPA axis function was assessed by 24-h urinary free cortisol (24-h UFC) excretion. Inflammation was assessed by high-sensitive C-reactive protein (hs-CRP). Multivariable regression analyses revealed a negative association between the lifetime score of adverse life events and HRV-HF β = -0.028; p = 0.033, but not with 24-h UFC or hs-CRP. We found limited support for the theory of allostatic load in a large population based cohort, with the exception of a modest association between ANS function and adverse life events.

INTRODUCTION

Psychosocial stress is a well known risk factor for cardiovascular disease (Rozanski et al. 1999, Ohlin et al. 2004, Rosengren et al. 2004). Explanatory models of how stress affects the body frequently assume that there are physiological pathways through which stressors can affect health. The detrimental effects of psychosocial stress on health in these models are usually attributed to a dysregulation of the stress responsive systems (Chrousos & Gold 1992, McEwen 1998b). The three major stress responsive systems postulated to be involved are the autonomic nervous system (ANS), the hypothalamic-pituitary-adrenal axis (HPA axis), and the immune system (Chrousos & Gold 1992, McEwen 1998a). Activation of these systems is thought to be adaptive in the short run. However, if these systems are activated too frequently, or if they fail to shut off after stress exposure, this activation might in the long term result in damage. It is this cumulative cost of stress exposure ('wear and tear') that is referred to as 'allostatic load' (McEwen & Stellar 1993).

Many studies have set out to test the idea that stress-responsive system dysregulation is associated with adverse cardiovascular health outcomes. Regarding the ANS, lower heart rate variability in the high frequency range (HRV-HF), signifying vagal withdrawal, has been shown to be an independent predictor of cardiac events in both healthy individuals and patients with a history of myocardial infarction (Bigger et al. 1993). Regarding the HPA axis, the cortisol-testosterone ratio is prospectively associated with incident ischemic heart disease (Smith et al. 2005). Regarding the immune system, recent research has demonstrated that minor elevations of C-reactive protein (hs-CRP) are a positive diagnostic predictor for future major cardiovascular events (Kushner et al. 2006).

However, as well-established as the connection between the abovementioned parameters of dysregulation of the stress responsive systems and cardiovascular disease might be; the connection between lifetime psychosocial stress and prospective dysregulation of the stress responsive systems in the general population is not. The results of studies focusing on the effects of stressors (e.g. adverse life events or chronic difficulties) on the function of the stress responsive systems are often conflicting. Altered activity of the ANS under resting conditions has been shown in individuals exposed to chronic stress compared to controls (Lucini et al. 2007, Lucini et al. 2008). In contrast, no differences in ANS activity have been shown in individuals with a history of adverse life events compared to controls (Otte et al. 2005, Pole et al. 2007). A meta-analysis on the association between stress exposure and HPA axis activity reported the daily output of cortisol to be high if the stressor was still present, but low if the stressor had been present in the past (Miller et al. 2007). A meta-analysis on the association between stressors and the immune system revealed different effects of different

types of stressors. Chronic stress may cause a general suppression of immunity, whereas non-specific life events may lead to suppression only in the elderly (Segerstrom & Miller 2004). The interpretation of previous studies is limited by three methodological shortcomings. First, they generally have a small sample size and ere often performed in specific subgroups (e.g. white-collar workers (Lucini et al. 2007) or students (Lucini et al. 2002)). Although the meta-analyses did include sufficient numbers of subjects, the generalizability of their results is questionable because the summary estimates were based on combinations of small case-control studies. Second, previous studies used heterogeneous measures of stress, for example including one specific type of stressor (e.g., childhood maltreatment (Danese et al. 2008) or caregivers stress (Lucini et al. 2008)) or stress exposure over a short time span (Lucini et al. 2002). Consequently, they cannot provide an insight in the cumulative effects of lifetime stress exposure, (i.e., whether psychosocial stress truly causes 'wear and tear' of the stress responsive systems). The only study that used an integrated measure was performed in elderly Taiwanese participants, thus suffering from the first mentioned problem (Glei et al. 2007). Third, previous studies often failed to control for relevant confounders or moderators, such as smoking (Kushner et al. 2006, Badrick et al. 2007), depressive disorder (Burke et al. 2005, Kushner et al. 2006, Rottenberg 2007), frequency of exercise (Britton et al. 2007), body mass index (BMI) (Ukkola et al. 2001), and antihypertensive medication (Schroeder et al. 2003). For these reasons, at present, it remains unclear whether cumulative psychosocial stress is associated with a dysregulation of the stress responsive systems in the general population.

The aim of the present study is to test the cumulative effect of stress on stress-responsive systems in a large population based cohort while taking into account known confounders and moderators. Based on the allostatic load theory, we hypothesized that adverse life events are negatively associated with HRV-HF, and positively associated with 24-hour urinary free cortisol (24-h UFC) and hs-CRP levels.

METHODS

Study population

Our study has been performed in a cohort derived from Prevention of REnal and Vascular ENd stage Disease (PREVEND), a population cohort study originally designed to investigate microalbuminuria as a risk factor for renal and cardiovascular disease. The recruitment of participants for PREVEND has been extensively described elsewhere (Pinto-Sietsma et al. 2000). All inhabitants of the city of Groningen between the ages of 28 and 75 years (85421 subjects) were asked to send in a morning urine sample and to fill out a short questionnaire on

demographics and cardiovascular history. A total of 40856 subjects (47.8%) responded. After exclusion of subjects with insulin dependent diabetes mellitus and pregnant women, all subjects with an elevated urinary albumin concentration of ≥10 mg/l (*N* = 7768) together with a randomly selected control group with a urinary albumin concentration of <10 mg/l (*N* = 3395) were invited for further investigations (total *N* = 11163). Finally, 8592 subjects completed the total screening program, rendering the PREVEND study cohort. Because the PREVEND study population was enriched for albuminuria, this oversampling for albuminuria was rectified in the current sub study. Albuminuria-negative participants were combined with a random sample of albuminuria-positive participants until a population-representative ratio was achieved. Research assistants handed over invitations to 2554 subjects to participate in the current sub study for which additional psychiatric and psychosocial data were collected. Measurements were completed by a total of 1094 participants (43%), forming our study cohort. PREVEND participants who declined to participate in the current study did not significantly differ from those who did participate concerning gender, age, and scores on a 12-item neuroticism scale (Sanderman et al. 1991). The study was approved by the medical ethics committee and conducted in accordance with the guidelines of the declaration of Helsinki. All participants gave a written informed consent.

Assessment of psychosocial stress

Participants completed the List of Threatening Experiences (LTE) at home prior to their visit of our research facilities. At the research facility a researcher checked with the patient if the list had been filled out correctly. The original LTE comprises 12 items, which are listed in Table 1, that were selected for their established long term consequences (Brugha et al. 1985, Brugha & Cragg 1990). In the original LTE, it is asked whether or not these events took place in the previous year. For the purpose of this study, the LTE was extended with the addition of age categories. In addition to the question about the occurrence of adverse life events in the previous year, the altered questionnaires comprise questions about the occurrence of life events in the following age categories: 0 - 12 years, 13 - 18 years, 19 - 39 years, 40 - 60 years, and >60 years. For each age category, the participants were asked to indicate whether or not the event had occurred. A sum score of the LTE (SUM-LTE) for each participant was calculated by adding the scores of all the different life events of all the age-categories (with 5 age categories and 12 life events the maximum score is 60). All participants filled in the LTE again approximately two years after the first assessment of these questionnaires. The test-retest reliability of the SUM-LTE with a two year interval was Pearson's r = 0.669. In this correlation coefficient, only age categories that were already completed at the first assessment were included.

Table 1. *The list of threatening experiences.*

Life event
- You yourself suffered a serious illness, injury or an assault
- A serious illness, injury or assault happened to a close relative
- Your parent, child or spouse died
- A close family friend or another relative (aunt, cousin, grandparent) died
- You had a separation due to marital difficulties
- You broke off a steady relationship
- You had a serious problem with a close friend, neighbour or relative
-You became unemployed or you were seeking work unsuccessfully for more than one month
- You were sacked from your job
- You had major financial crisis
-You had problems with the police and a court appearance
- Something you valued was lost or stolen

Assessment of HRV-HF

Analysis of beat-to-beat variations in heart rate (HR; i.e., heart rate variability, HRV) provides a non-invasive method to measure autonomic function. At both baseline and two-year follow-up, HRV measurements were performed in the afternoon, with participants lying on a bench in the supine position in a quiet laboratory room, breathing spontaneously. Research assistants who measured HRV were blind to our study hypotheses. There were no restrictions in eating, drinking, or smoking in the hours prior to the measurement. Participants were encouraged to relax and asked not to move or speak during data acquisition. A pre-rest period of 10 minutes in the supine position was applied before the HRV measurement started. A cuff was fixed around the middle phalanx of the third finger on the right hand. A Portapres device (FMS Finapres Medical Systems BV, Amsterdam, the Netherlands) continuously recorded HR during 15 minutes. Segments with a duration of 300 seconds were selected for spectral analysis. In case there was no appropriate segment of 300 seconds, blocks of 60 - 300 seconds were selected. Power spectral analysis of HR signals was performed with the CARSPAN software program (Mulder 1988). The measurements were checked on stationarity and corrected for artifacts. Artifacts were eliminated and the resulting gaps were linearly interpolated. When a data set contained more than 10% interpolated HR intervals or too many artifacts, the data were considered unstable and discarded. Participants for whom no reliable HRV measurement was available (<5% at both measurements waves) were comparable to participants with reliable HRV measurements regarding gender and age. The high frequency band (HRV-HF), defined at 0.15 – 0.40 Hz, is expressed in ms^2 and mainly reflects parasympathetic activity (Task Force of the European Society of

Cardiology and the North American Society of Pacing and Electrophysiology 1996, Berntson et al. 1997). After ln transformation, HRV-HF values are expressed in ln (ms^2).

Assessment of 24-h UFC

Participants were asked to collect urine samples in a polypropylene container on two consecutive days prior to the visit to the outpatient clinic. They were instructed to urinate into the container during the 24-h collection period and refrigerate the sample until delivery to the laboratory. UFC was measured by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Use of LC-MS/MS is recommended because this method offers advantages over immunoassays methods; the LC-MS/MS method is free of interferences from cortisol metabolites and conjugates and also eliminates drug interferences (Taylor et al. 2002). 24-h UFC was calculated by multiplying urinary volume with cortisol concentration and is expressed in nmol per 24-h. We used the mean of the two samples on consecutive days to reflect HPA axis function. Additionally, urine creatinine concentrations (mmol/l) and body surface area (m^2) were assessed (DuBois & DuBois 1916). Repeating our analyses with 24-h UFC/urine creatinine ratio or 24-h UFC/body surface area ratio as dependent variable did not alter our results. Therefore, we used unadjusted 24-h UFC excretion in our analyses.

Assessment of hs-CRP

Fasting blood samples were collected in all participants during a visit to the research facilities. In case of flu or a febrile temperature, blood collection was postponed to a later time (available blood samples at baseline N = 1024, at follow-up N = 983). High-sensitive CRP (hs-CRP) was determined by nephelometry with a threshold of 0.18 mg/L and intra- and inter-assay coefficients of less than 4.4 and 5.7%, respectively (BNIIN, Dade Behring, Marburg, Germany). CRP levels below the detection level were scored as 0.18 mg/L. One-time measurement is a reliable measure of CRP over an extended period of time in healthy individuals (Macy et al. 1997). Since concentrations of CRP increase dramatically as part of acute inflammatory response, it has been recommended to discard plasma CRP concentrations above 10 mg/L in studies on CRP as a risk factor for cardiovascular disease (Pearson et al. 2003). Participants with a score of 10 mg/L and above (N = 44) were therefore excluded from analyses with hs-CRP as outcome variable.

Statistical analysis

Multivariable regression analyses were performed with as forced entered predictor variables the measure of psychosocial stress (i.e., SUM-LTE), gender, and age. In all cases, the outcome variable was one of the measures of the three stress responsive systems (i.e., HRV-HF, 24-h UFC, or hs-CRP). All outcome variables were checked for normality of their distribution and log transformed to meet this

assumption when necessary. Data were checked for outliers, +/- 3 standard deviations, if assumptions of a normal distribution had been met. In the case of 24-h UFC, which had a non-Gaussian distribution, outliers were detected with the Hampel identifier (Wilcox 2002). Regression analyses were adjusted for the following potential confounders and moderators: gender, age, BMI, presence of depressive disorder (current diagnosis of major depression as assessed by the Composite International Diagnostic Interview), smoking (none, 1-5 cigarette(s)/day, 6-10 cigarettes/day, 11-15 cigarettes/day, 16-20 cigarettes/day, >20 cigarettes/day), and frequency of exercise (not/hardly, once per week, twice or more per week). In the ANS analyses, we excluded participants on antihypertensive medication and adjusted for the use of antidepressants. In the HPA axis analyses, we excluded participants on corticosteroids and adjusted for the use of antidepressants. In the hs-CRP analyses, we excluded participants on antihypertensive medication, corticosteroids, and oral contraceptives, and we adjusted for analgesics use. We also present results of regression models that are only adjusted for gender and age to illustrate the influence of the covariates. Results were considered statistically significant for a p-value <0.05. Data were analyzed using the Statistical Package for Social Sciences version 14.0 (SPSS Inc, Chicago, IL, USA).

RESULTS

Descriptive statistics about the characteristics of our study population are provided in Table 2. The results of the multivariable regression analyses showed that SUM-LTE was a significant negative predictor for HRV-HF (Table 3). On the contrary, SUM-LTE did not significantly predict 24-h UFC or hs-CRP (Table 3). After rerunning our analyses adjusting for lifestyle factors and medication, the results remained essentially unaltered (Table 4). In addition, post-hoc analyses showed no effect modification by gender or age (data not shown).

DISCUSSION

In a general population based cohort study, we found cumulative adverse life events to be a significant negative predictor of HRV-HF but not of 24-h UFC or hs-CRP. The results on the ANS thus supported the allostatic load theory, whereas the results on the HPA axis or the immune system did not.

There are several strengths and weaknesses of this study that need to be taken into consideration when interpreting our results. The first major strength of this study is that we conducted our study in a large population representative cohort.

Table 2. *General characteristics of the study population (N = 1094).*

	Female	**Male**
Gender (%)	53.7	46.3
Age (SD), range	52.5 (11.1), 33 - 79	53.8 (11.6), 33 - 78
Race (%)		
– Caucasian	97.4	
– Negroid	0.9	
– Asian	0.7	
– Other	0.3	
SUM-LTE, mean (SD)	5.7 (3.2)	4.9 (3.3)
HRV-HF ln (ms^2) median (IQ)	6.2 (5.4 - 6.9)	6.1 (5.3 – 6.8)
24-h UFC (nmol/24-h) median (IQ)	70.6 (43.8 – 111.9)	65.7 (42.3 – 99.5)
hsCRP (mg/l) median (IQ)	1.4 (0.6 – 3.2)	1.2 (0.6 – 2.7)
Antihypertensive medication (%)	12.2	13.4
Depressive disorder (%)	2.0	1.4
Antidepressants (%)	3.6	1.4
Smoking (yes) %	22.4	25.5
Exercise (%)		
– Not/hardly	49.8	53.8
– Once per week	28.7	26.1
– Twice or more per week	20.9	19.6
BMI (kg/m^2) mean (SD)	26,5 (4.5)	26.6 (3.6)
Corticosteroid use (%)	17.2	19.2
Oral contraceptives use (%)	10.2	n.a.

Abbreviations: BMI = body mass index, HRV-HF = high frequency band of heart rate variability, hsCRP = high sensitive C-reactive protein, IQ = interquartile range, ln = natural logarithm, ms^2 = square milliseconds, nmol/24-h = nanomol per 24 hours, mg/l = milligrams per litre, SD = standard deviation, SUM-LTE = lifetime score of adverse life events, 24-h UFC = 24-hour urinary free cortisol

Table 3. *Multivariable regression analyses predicting HRV-HF, 24-h UFC, and hs-CRP by SUM-LTE adjusted for gender and age.*

	HRV-HF ***N* = 881**				**24-h UFC** ***N* = 986**				**hs-CRP** ***N* = 967**			
	r^2	β	*t*	*B*	r^2	β	*t*	*B*	r^2	β	*t*	*B*
SUM-LTE	.061	-.028*	-2.142	-0.069	.002	.001	.387	.012	.042	-.005	-1.185	-.038

*Abbreviations: HRV-HF = high frequency band of heart rate variability, hs-CRP = high-sensitive C-reactive protein, SUM-LTE = lifetime score of adverse life events, 24-h UFC= 24-hour free urinary cortisol. *p = 0.033.*

Table 4. *Multivariable regression analyses HRV-HF, 24-h UFC, and hs-CRP by SUM-LTE, fully adjusted for confounders and moderators.*

	HRV-HF † ***N* = 774**				**24-h UFC ↕** ***N* = 731**				**hs-CRP ‡** ***N* = 613**			
	r^2	β	*t*	*B*	r^2	β	*t*	*B*	r^2	β	*t*	*B*
SUM-LTE	.076	-.028	-1.920	-.068	.016	.003	.967	0.036	.248	-.006	-1.149	-.042

Abbreviations: hs-CRP = high-sensitive C-reactive protein, HRV-HF = high frequency band of heart rate variability, SUM-LTE = lifetime score of adverse life events, 24-h UFC = 24-hour urinary free cortisol, † = antihypertensive medication excluded and adjusted for gender, age, depressive disorder, antidepressants, smoking, exercise, and BMI, ↕ = corticosteroids excluded and adjusted for gender, age, depressive disorder, antidepressants, smoking, exercise, and BMI, ‡ = antihypertensive medication, oral contraceptives, and corticosteroids excluded and adjusted for gender, age, depressive disorder, analgesics, smoking, exercise, and BMI.

Secondly, a strength of our measure of psychosocial stress is that we measured it over the entire lifespan. This makes it possible to look at the additive effects of stress over life, whereas other studies measured stressful events usually only in the previous year.

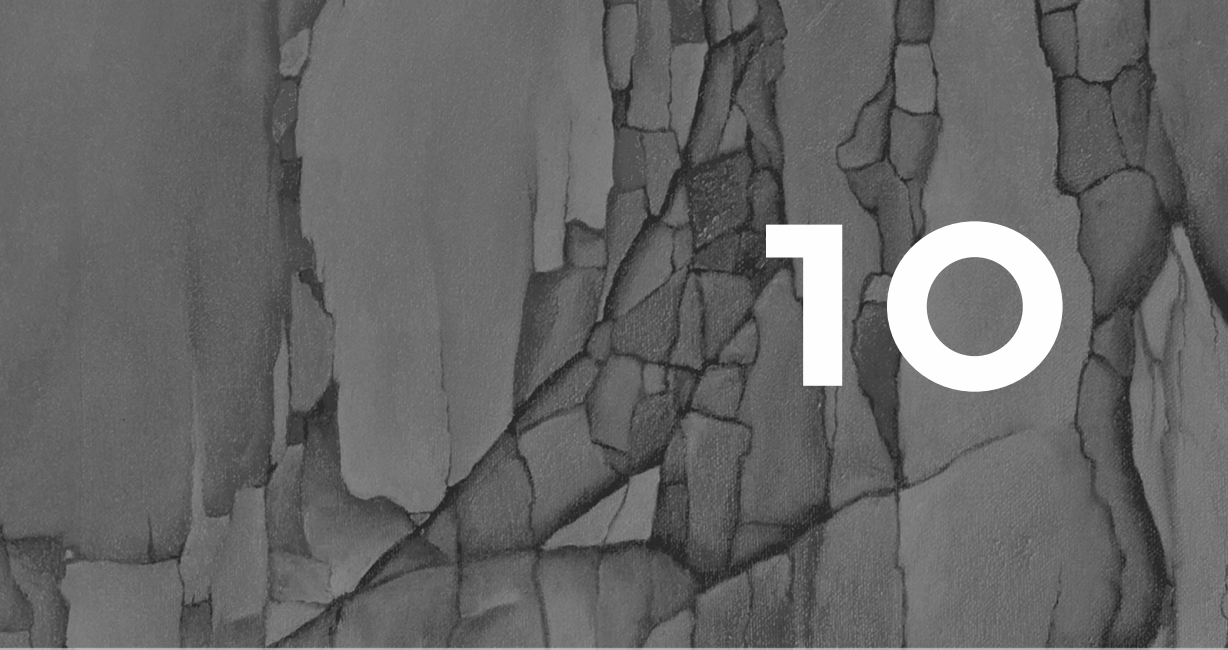
The most important limitation of this study is that we measured only one component of each stress responsive system. An important drawback of using 24-h UFC as a measure of HPA axis function is that it does not give an insight in possible changes in the diurnal rhythm of cortisol excretion. It is nevertheless an integrated measure of HPA axis function not influenced by the time of measurement and interpersonal differences in circadian rhythm. Besides, a large meta-analysis on psychosocial stress and HPA axis function shows that, although different stressor types have different impacts on the circadian profile, the total 24-hour cortisol output was usually also significantly altered (Miller et al. 2007). Regarding HRV, we did not measure respiratory rate and depth, whereas it has been recommended to monitor and adjust for these characteristics. However, this requirement seems particularly relevant when studying within-subject changes in HRV (Ritz & Dahme 2006). Correction or control procedures are discouraged in between-subjects designs such as in our study, since HRV-HF appears not dependent on respiration frequency under baseline conditions (Denver et al. 2007). Finally, although we have demonstrated that test-retest reliability was adequate, we cannot exclude the possibility of recall bias due to the use of retrospective questionnaires (Streiner & Norman 2003). In addition, as participants only indicated whether the life event had occurred and not how many times it occurred in each age category, we might have underestimated the effect of psychosocial stress on the stress responsive systems.

Our results for the ANS support the theory of allostatic load. Previous studies from one other group reported no differences in ANS activity in individuals with a history of adverse life events compared to controls (Otte et al. 2005, Pole et al. 2007). However, these authors specifically studied police academy recruits, using catecholamine and startle responses as measures for ANS activity. Our results are in agreement previous studies that found a negative association between chronic psychosocial stress (e.g., job strain, cancer care-givers stress) and HRV-HF (Collins et al. 2005, Lucini et al. 2007, Lucini et al. 2008).

In contrast to the results on the ANS, our results on the HPA axis or the immune system did not support the allostatic load theory, in contrast to previous studies (Heim et al. 2000b, Dickerson & Kemeny 2004, Danese et al. 2007, Miller et al. 2007, Elzinga et al. 2008). This might be explained in a number of ways. First, most previous studies were done in selected and / or small samples, for example grouping patients (e.g., with depression) or subjects with a known history of adversities and comparing them to healthy controls. In contrast, previous studies investigating the ANS were conducted in healthy participants and were in this

respect more similar to our study. Second, we used a composite score based on different types of stressors, whereas most of the studies investigated the effects of highly traumatic events. It can be debated whether such high impact events have the same effect on the function of the HPA axis and hs-CRP as less disruptive ones. Maybe the effects of psychosocial stress on the HPA axis and hs-CRP were too small to be detected or the effects on the ANS were more robust. Third, most studies investigating the effects of psychosocial stress on the HPA axis (Miller et al. 2007) or the immune system (Segerstrom & Miller 2004) do not appropriately adjust for confounding factors. The authors of these meta-analyses rightfully comment on this issue in their recommendations for future research. Interestingly, most studies in the body of literature on the ANS did appropriately adjust for confounders (Lucini et al. 2005, Collins et al. 2005, Lucini et al. 2007, Lucini et al. 2008), suggesting that the association might be genuine.

Psychosocial stress is a well known risk factor for cardiovascular disease (Rozanski et al. 1999, Rosengren et al. 2004). In addition, lower HRV has been shown to be an independent predictor of cardiac events (Bigger et al. 1993, Tsuji et al. 1996). In the present study, in conformity with the theory of allostatic load (McEwen & Stellar 1993), it is demonstrated for the first time that cumulative psychosocial stress is associated with deteriorated functioning of the ANS. Therefore, the physiological consequences of exposure to adverse life events might in part be responsible for the association between psychosocial stress and cardiovascular disease. Prospective research is needed to not only unravel the relationship between psychosocial stress and the stress responsive systems, but also to disentangle the relationship between psychosocial stress and disease outcomes with dysregulation of the stress responsive systems or stress altered health behavior as possible mechanistic links. Finally, future studies investigating the HPA axis or the immune system should adequately control for confounders, which at present, unfortunately, is still a rarity.



10

CHAPTER 10

General discussion:
Is there a pivotal role for dysfunction of stress responsive systems in the etiology of functional somatic symptoms and disorders?

LM Tak and JGM Rosmalen
Journal of Psychosomatic Research, in press

BACKGROUND

The etiology of somatization, a collection term for the processes involved in experiencing functional somatic symptoms (FSS) and disorders (FSD), is generally considered to be a multifactorial interplay between psychological, biological, and social factors (Deary et al. 2007). The idea of stress responsive system dysfunction as a mediator between psychosocial stress and somatization has often been advanced as one of the potential biological factors. Although intuitively attractive, more than twenty years of research has not made clear whether and how dysfunction of stress responsive systems is involved in the etiology of somatization. Therefore, the aim of this thesis was to investigate the role of dysfunction of stress responsive systems in the etiology of FSS and FSD.

It has been assumed that dysfunction of stress responsive systems, consisting of alterations in activity of the autonomic nervous system (ANS), hypothalamic-pituitary-adrenal (HPA) axis, and immune system, contributes to somatization. This assumption is largely based on cross-sectional case-control studies in the main three functional FSD, namely, chronic fatigue syndrome (CFS), fibromyalgia (FM), and irritable bowel syndrome (IBS). Although FSD are usually researched in isolation, they share many factors (Wessely et al. 1999). Those shared factors might underlie susceptibility for development of any FSD, whereas FSD-specific factors might shape their final manifestation (Aggarwal et al. 2006, Kato et al. 2009). As outlined in the introduction of this thesis, findings of the studies on dysfunction of stress responsive systems in FSD were inconclusive. They should be interpreted with caution because of the general problems with evidence derived from case-control studies, such as selection bias and confounding. Moreover, like many psychosomatic studies, they have characteristics that contribute to the risk of non-replication, such as having small sample sizes and large flexibility in designs, definitions, and outcomes (Ioannidis 2005, Freedland et al. 2009). Small sample sizes will have reduced power to detect true differences or relevant subgroups (i.e., a small overall difference in mean values between cases and controls may have been caused by a large effect in a subgroup of cases with no effect in the remainder). Finally, the cross-sectional design of those studies only allowed inferences on whether stress responsive dysfunction is a correlate but not whether it is a risk factor for somatization (Kraemer et al. 1997). In this thesis, we have adopted two strategies to overcome these problems. First, in Chapter 2, we argued that meta-analyses can contribute to more reliable knowledge about the existence and summary estimate of certain associations in the area of psychosomatic research. Apart from providing this summary estimate, the importance of meta-analysis also lies in the identification of reasons behind contradictory findings and the construct of guidelines to conduct better quality research in the future. Therefore, we performed a critical synthesis of previous research findings on ANS and HPA axis activity in three main FSD (Chapters 3

and 4). Regarding immune function in FSD, there were not enough eligible studies to allow quantitative analysis. A second solution to avoid the methodological problems inherent to case-control studies was conducting adequately powered population-based studies that take into account important factors as detected in the meta-analyses. We performed population-based studies on the association between the three stress responsive systems and FSS (Chapters 6 - 8). Chapter 5 provides the basis for using the total number of FSS as a measure of somatization. Previously suggested bodily clusters fitted our data (i.e., cardiopulmonary, musculoskeletal, gastrointestinal, and general FSS clusters (Fink et al. 2007)), but the diagnosis of somatization could be based on a simple cut-off of four FSS, regardless of the number of symptom clusters from which they originate.

Previous research has usually been of cross-sectional nature. We have introduced different pathways to summarize hypothetical relations among stress responsive system dysfunction and somatization: the etiological pathway, in which alterations in stress responsive system function induce FSS and FSD; the epiphenomenal pathway, in which stress responsive system dysfunction and FSS and FSD share the same etiological determinants but are not causally related; and the consequential pathway, in which FSS and FSD induce stress responsive system alterations that are either causally unrelated or perpetuate symptom experience. Finally, given the contradictory available evidence, the possibility that there is no association between stress responsive system dysfunction and FSS and FSD should also be considered. Because stress responsive system activity and FSS were measured at two time points in our population-based cohort, we were able to gain more insight in temporal associations.

The aim of this discussion is to summarize the results and interpret the current evidence for a pivotal role of stress responsive system dysfunction in FSS and FSD. We will integrate the studies of this thesis as well as other relevant research, such as prospective research and research focusing on symptom experience. For each stress responsive system, we will give an overview of available evidence as to whether alterations can be considered as a causal risk factor, variable marker, fixed marker, concomitant or consequence, correlate, or noncorrelate of somatization (Kraemer et al. 1997). We will also discuss whether alterations are shared or FSD-specific. We will comment on methodological problems in the field and provide solutions where possible. Many times the methodological issues outlined for one the three stress responsive systems theoretically applies to all. Finally, we conclude with answering the question whether there is a pivotal role of stress responsive system dysfunction in somatization, give directions on how to successfully proceed and ultimately benefit patients by improving our understanding of this topic.

ANS ACTIVITY AND SOMATIZATION

Are alterations in ANS activity correlated with somatization?
Our meta-analysis of ANS activity in three different FSD indicated statistically significant lower baseline cardiac vagal activity as measured by heart rate variability (HRV) in FSD patients compared to controls (Chapter 3), with no apparent differences between CFS, FM, and IBS. The validity of these summary estimates was, however, significantly limited by unexplained heterogeneity in the effect sizes of included studies. Unfortunately, the number of included studies was too small to allow for moderator analyses that may detect variables explaining this observed heterogeneity in effect sizes, such as differences in gender, medication use, psychiatric co-morbidity, or physical inactivity. This meta-analysis also showed that there is substantial room for improvement in recruitment of participants, especially in selection of healthy controls. Instead of consisting of students or staff members, a research topic that relies on subjective measures (i.e., self-reported symptoms) and measures that are related to several conditions (i.e., alterations in ANS activity) strictly requires that healthy controls come from the same population as cases. In order to recruit a representative sample of healthy controls, nested case-control designs in population-based studies seems the optimal strategy. Blinding of researchers who perform HRV measurements, report of adequate HRV outcomes, and adjustment for potential confounders emerged as other important methodological factors that need improvement. Furthermore, when taking the presence of potential publication bias into account, the difference in cardiac vagal activity between FSD patients and controls disappeared, suggesting that some studies contrasting with the prevailing beliefs have not been published. A population-based study that was published afterwards, and adjusted for factors such as gender, age, body mass index, medication use, and physical inactivity, found no difference in cardiac vagal activity between CFS subjects and controls (Boneva et al. 2007). Another recent study in patients with somatization disorder found no differences in cardiac vagal activity compared to controls (Laederach-Hofmann et al. 2008), but concluded that 62% of the patients had autonomic dysregulation based on baroreflex sensitivity, another measure of ANS function. We felt, however, that this firm conclusion was not justified by the analyses as presented (Tak et al. 2009c), illustrating the clear need for methodological sound studies in this field.

Our own population-based study on ANS activity in adults with FSS demonstrated that decreased cardiac vagal activity is associated with a higher number of FSS in young adults ≤52 years. Moreover, as was the case in the meta-analysis, this association with lower cardiac vagal activity was generic in the sense that is was similar for different bodily clusters of FSS (e.g., musculoskeletal cluster resembling FM and the gastrointestinal cluster resembling IBS). Although this thesis focuses on adults, it is interesting to mention that a large population-based study of 920 pre-adolescents also found that cardiac vagal activity was also

negatively related to FSS, in both boys and girls (personal communication Dr. A. Dietrich). In contrast, higher cardiac vagal activity is associated with a higher number of FSS in older adults >52 years, a finding that was unexpected. Differential etiology of FSS in older compared to younger persons may underlie this unexpected interaction effect, however, this does in itself not explain the higher cardiac vagal activity. Cardiovascular co-morbidity (Thayer & Lane 2007) or hypertension-associated hypoalgesia (Ghione 1996) did not explain this association. A promising, but largely unexplored hypothesis is that high HRV reflects different processes in younger persons than in older persons, due to age-related changes in cardiovascular functioning (Stein et al. 2005, De Meersman & Stein 2007). An unanticipated interaction between age and cardiac vagal activity in psychophysiological research has been shown before in a study on depressive disorder. While cardiac vagal activity generally decreases with increasing age (De Meersman & Stein 2007), it has been observed that cardiac vagal activity remained stable with increasing age in depressed patients but not in controls (Jindal et al. 2008). In this study, inspection of regression lines suggests that between-group differences might be evident before age 60, but not thereafter. This cut off seems to be in line with the age-associated differences in the association between cardiac vagal activity and FSS in our population-based study (Chapter 6).

Do alterations in ANS activity precede somatization?

Available longitudinal evidence on whether alterations in ANS activity precede somatization is scarce. In our population-based study on ANS activity and somatization, decreased cardiac vagal activity is associated with FSS after two years follow-up in younger adults. Increased cardiac vagal activity is unexpectedly associated with FSS after two year follow-up in middle-aged to older adults. Notably, this is the same pattern as in the cross-sectional part of the study; however, despite a significant interaction term, stratified analyses did no longer reach statistical significance after two year follow-up (Chapter 3). We are aware of one other long-term longitudinal ANS study in FSS, concluding that spontaneous improvement of FSS in a non-clinical sample could not be predicted by ambulatory physiological recordings of cardiac vagal activity a year earlier (Houtveen & van Doornen 2008). No longitudinal studies that explore the issue whether ANS alterations are associated with the development of FSD have been published. We are aware of one study that has assessed whether ANS activity, as measured with HRV, could be directly related to symptom experience. This study was performed in a non-clinical student population of 18 young females scoring high on a list of hyperventilation-related FSS and 18 young females scoring low on this list. Although the group high on FSS reported a significantly larger number of somatic symptoms after mental stress and breathing of CO_2-enriched air compared to the group low on FSS, there were no accompanying differences in cardiac vagal activity (Houtveen et al. 2003). The authors argue that persons

scoring high on hyperventilation-related FSS possibly have an exaggerated perception of normal peripheral physiology.

Can ANS activity change or be changed?

In contrast to a fixed marker, a variable risk factor is a risk factor that can be demonstrated to change spontaneously within a subject or to be changed within a subject by intervention (Kraemer et al. 1997). The ANS can be altered by several factors, such as age (De Meersman & Stein 2007), medication use (Rechlin 1994, Agelink et al. 2002, Schroeder et al. 2003), and psychosocial stress (Chapter 9). In Chapter 9, our objective was to assess whether self-reported adverse life-events during the lifespan are associated with current activity of stress responsive systems using data measured in the population-based cohort. The only stress responsive system that was associated with life events was the ANS, which is consistent with what we would expect if a causal path leads from chronic psychosocial stress to alterations in ANS activity to somatization. Regarding the capacity to change, there is some preliminary evidence that interventions, such as HRV biofeedback (Nolan et al. 2005) or meditation (Tang et al. 2009) can improve cardiac vagal activity. It should be noted that ANS alterations are aspecific, and have been associated with several somatic and psychiatric conditions (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996, Kuch et al. 2001, Haapaniemi et al. 2001, Schroeder et al. 2003, Licht et al. 2008). It has, for example, been documented that the associations between HRV and depressive and anxiety disorders are mainly driven by the effects of antidepressants (Licht et al. 2008, Licht et al. 2009). However, in our study we performed sensitivity analyses after taking antidepressant medication use into account, and the associations with somatization remained (Chapter 6). Although guidelines on measurement of HRV are available (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology 1996), general recommendations about covariates to consider as confounder, mediator, or moderator are lacking. Covariates should be a priori chosen based on theory, as this strategy will generally constitute stronger scientific evidence than statistical models that are achieved in other ways (Babyak 2004). However, understanding of the role of covariates in the association between dysfunction of stress responsive systems and somatization is still in its infancy. Consequently, it is often difficult to decide how to consider covariates. Not adjusting for certain covariates may lead to spurious associations, whereas overadjusting may lead to attenuation of genuine associations. In Chapter 3, we provide a quality tool with recommendations for exclusion and control variables when studying the association between the ANS and FSD. Still, conceptual problems remain. As an example, consider the role of physical inactivity. We have adjusted for physical inactivity in our population-based studies because we considered this variable as a confounder (Chapter 6). This strategy may be too conservative in case physical

inactivity is a causal risk factor of symptom experience with stress responsive system dysfunction as mediator (Christenfeld et al. 2004). Furthermore, the association between ANS activity and somatization may differ by gender or age. In other words, gender and age may be moderators. A moderator is a covariate that affects the direction and/or strength of the association between the dependent and independent variable (Rothman 1976, Baron & Kenny 1986). Our study on ANS function and FSS revealed that age might be a moderator, as different values of cardiac vagal activity were found in younger compared to older adults (Chapter 6). Interestingly, the extent to which ANS activity can be changed (and thus, the extent to which it is a variable risk factor instead of a fixed marker) may also differ by gender or age. For example, exercise training results in significant increases in cardiac vagal activity, but these effects were considerably larger in younger compared to older subject (Sandercock et al. 2005a). Gender also has a significant role in these exercise-related changes in healthy young adults, as cardiac vagal activity increased after training and decreased after deconditioning in men, whereas no significant changes were apparent in women (Sloan et al. 2009).

Does manipulation of ANS activity change somatization?

No study assessed whether cardiac vagal activity can be manipulated to change somatization outcome. However, some circumstantial arguments may plead for a causal pathway, such as the finding that the ANS is the only stress responsive system measure that was associated with chronic psychosocial stress (Chapter 9) and with having a stress responsive personality, i.e., negative affectivity or neuroticism (Bleil et al. 2008). The argument pleading against causality is the finding that change in symptom experience was not accompanied by alterations in cardiac vagal activity (Houtveen et al. 2003). However, this finding in a specific subgroup of FSS in healthy students should be replicated in larger populations suffering from clinically relevant somatization. In future research, innovative designs could be used in persons with FSS or FSD, such as simultaneously measuring HRV as well as objective and subjective measures of pain (Paine et al. 2009).

Conclusions on the role of ANS activity in somatization

Given the potential publication bias, the overall poor methodological quality, and the lack of longitudinal studies, it seems not possible to firmly reject or accept a role of ANS dysfunction in FSD or to differentiate about its relative importance across the three FSD. However, our own population-based study points in the direction that lower cardiac vagal activity is a shared risk factor for somatization in younger adults. The observation that lower cardiac vagal activity precedes new-onset FSS suggests that alterations in this stress responsive system are not just a concomitant or consequence of FSS. The intriguing finding of higher cardiac vagal activity in older adults with FSS warrants further research and better understanding. Based on current available evidence, it seems even premature to

decide whether ANS activity is a correlate of somatization, let alone whether it is a causal risk factor. Overall, evidence suggests reduced parasympathetic activity as a generic factor, but the inconclusive evidence on the role of the ANS in FSS and FSD requires further study.

HPA AXIS ACTIVITY AND SOMATIZATION

Are alterations in HPA axis activity correlated with somatization?

As summarized in the introduction, findings of more than sixty cross-sectional case-control studies in various FSD did not exclusively support one separate kind of dysfunction of the HPA axis. The findings of the association between the HPA axis and FSD included mild hypocortisolism, but normal cortisol levels, as well as hypercortisolism are also reported (Tak & Rosmalen 2007). Reviews specifically focusing on patient with CFS or FM have concluded that cortisol levels, when altered, tend to be in the lower part of the normal range (Geenen et al. 2002, Cleare 2003, Tanriverdi et al. 2007). As heterogeneity of various confounding, mediating, and moderating variables may have contributed to the divergent findings to date (Cleare 2003), it has been recommended to measure those variables more precisely to allow a better insight in HPA axis alterations in FSD. The aims of our meta-analysis were therefore to assess whether FSD are characterized by HPA axis alterations and, if present, to examine by which variables this association was influenced (Chapter 4). We demonstrated that baseline cortisol levels are lower but not significantly different in FSD subjects compared to healthy controls by meta-analyzing 85 available cross-sectional studies. When FSD are considered separately, hypocortisolism is present in CFS and possibly in FM, but not in IBS. Lower cortisol levels in FSD subjects are predominantly found in studies which included a larger proportion of women, suggesting that the role of HPA axis activity is gender-specific. Several explanations can be offered to clarify why, in contrast to our hypothesis that it would be a shared factor, hypocortisolism is differentially related to FSD. CFS and FM may be etiologically more alike conditions due to shared hypocortisolism that contributes to the experience of fatigue and widespread pain, which are prominent features of CFS and FM, but not of IBS. Cortisol levels in IBS patient may also differ from those observed in CFS and FM because of the type of studies which are performed in IBS patients, which are often stressful procedures such as rectal extensions which may elicit acute anticipatory stress responses with a hyperactive HPA axis (Walter et al. 2006, Miller et al. 2007). A third explanation is that cortisol alterations may not be FSD-specific, but instead origin from certain subgroups among FSD, which have the highest prevalence in CFS. An example of a subgroup may be subjects with co-morbid depressive disorder, which indeed has a higher prevalence in CFS compared to IBS (Henningsen et al. 2003). Specifically, FSD patients might experience more atypical forms of depression

(American Psychiatric Association 1994), given that atypical depression is characterized by hypocortisolism (Gold et al. 1995, Antonijevic 2006). Supporting the importance of further elucidating the presence of atypical depression in somatization, a recent study found that a negative correlation between salivary free cortisol in the morning and depressive state was a characteristic of patients with heterogeneous FSD, but not of healthy controls, in whom cortisol and depressive state were positively correlated (Mutsuura et al. 2009). Finally, FSD-specific cortisol alterations may arise from differences in behavioral consequences of the FSD, such as changes in medication use, physical activity, sleeping pattern, or smoking habits. When studying the role of covariates, considering the quality of measurement is also important. For example, although studies often failed to measure level of physical activity, the ones who did usually relied on self-report. A study in CFS using actigraphy as an objective method to assess physical inactivity showed that only a subgroup of the CFS patients can be labeled as persistently inactive (van der Werf et al. 2000). When a covariate is not measured adequately, residual confounding or mediating may distort findings. A limitation of this meta-analysis was that some variables that might theoretically be important moderators could not be tested because they were seldom measured in the included studies. For example, acute psychosocial stress and psychosocial stress in the past are rarely addressed, but both impact cortisol levels (Dickerson & Kemeny 2004, Miller et al. 2007, Michaud et al. 2008). In this perspective, one relevant stressor might be childhood trauma, as it has recently been found that decreased cortisol responses to awakening are observed only in those individuals with CFS who reported exposure to childhood trauma but not in individuals without such exposure (Heim et al. 2009).

In agreement with the meta-analysis, an experimental study on 80 healthy young adults, those who scored higher on subjective measures such as pain intensity and pain unpleasantness, had a flattened cortisol awakening response (i.e., lower cortisol in the morning) as measured in the days after the pain assessment (Fabian et al. 2009). A small study that examined saliva cortisol samples of 14 healthy office workers during four consecutive weeks showed that low cortisol levels in the morning and high cortisol levels in the evening were associated with poor self-rated health and fatigue (Dahlgren et al. 2009). In our population-based study, taking the role of medication use into account and adjusting for a large range of potential confounders, there is no cross-sectional association between 24-h urinary free cortisol (24-h UFC) and the total number of FSS experienced in the previous year (Chapter 7). In addition, 24-h UFC excretion is not associated with the number of FSS in any of the bodily clusters. Notably, no interactions of gender or age were found in this study.

The influence of different decisions on how to measure then HPA axis activity (such as assessing cortisol in blood, saliva, or urine; taking samples in the morning, afternoon, evening, or during sleep; calculating absolute values at one

time point or the area under the curve) is largely unexplored. Single point measurements are generally considered not representative of HPA axis activity (Nicolson 2007). Many early studies of cortisol and health have focused on average cortisol measures, particularly urinary measures of cortisol. Increasingly, however, the focus shifts to the marked diurnal rhythm in the release of cortisol (Adam & Kumari 2009). The diurnal cortisol rhythm is typically characterized by high levels upon waking, a substantial (50 - 70%) increase in cortisol concentration in the 30 - 45 minutes after wakening (i.e., the cortisol awakening response), and a subsequent decline over the remainder of the day, reaching a low point or nadir around midnight (Pruessner et al. 1997, Edwards et al. 2001). In a meta-analysis, the cortisol awakening response has been shown to be positively associated with job stress and general life stress, and negatively associated with fatigue, burnout, and exhaustion (Chida & Steptoe 2009). In this perspective, the observation that cross-sectional cortisol differences between FSD subjects and healthy controls are especially apparent in morning samples is interesting (Chapter 4). To obviate some of the differences in measurements, recommendations to harmonize saliva collection protocols in large-scale survey settings have been made (Adam & Kumari 2009). For example, it is recommended that population-based research should involve one sample collected on waking, one at the peak of the cortisol awakening response (which typically is 30 minutes after waking) and one at bedtime. Addition of more days of data provides more important information than addition of more data points within a single sampling day. In combination with the quality requirements we have proposed in Chapter 4, future studies on stress responsive system and somatization should adhere to those guidelines.

Although we have specifically focused on baseline measurements, HPA axis activity can also be measured after a challenge or stress test. The rationale behind stimulating the HPA axis is that the level of the underlying pathophysiology (i.e., hypothalamus, pituitary, or adrenals) and more subtle alterations may become visible. In subgroups of somatization, challenge or stress tests may be particularly interesting, for example IBS patients in whom acute stress is thought to play a role (Patacchioli et al. 2001). On the other hand, acute stress does not seem to evoke FSS (Burton et al. 2009) and symptoms are usually present at a larger part of the day and not only in stressful situations. Another problem is the variety of applied measurements that is even larger compared to baseline measurements, as many different pharmacological and non-pharmacological stressors are used. Psychosocial stress tasks are often thought to have the greatest ecological validity in this type of research (Nicolson 2007). Among the psychosocial stressors, performance tasks with elements of social-evaluative threat or uncontrollability produce the largest and most consistent increases in cortisol (Dickerson & Kemeny 2004). Only a few stress tasks have been described in sufficient detail that results can be compared across studies and populations, such as the Trier

Social Stress Test (Kirschbaum et al. 1993). However, it can also be speculated that standardization is artificial, and gaining insight in pathophysiology of symptom experience needs to be tailor made stress tests, such as asking participants to do something that they expect to elicit symptoms. It is also important to realize that rest and stress measurements within a person appear poorly interrelated. For example, the cortisol awakening response and cortisol values after a psychosocial stress test are not correlated (Bouma et al. 2009). This is surprising given the general idea that they are representatives of the same underlying HPA axis pathology. Thus, measuring cortisol in rest may reveal other underlying mechanisms than measuring cortisol after challenge or stress tests. Even the naturalistic cortisol awakening response and cortisol levels later in the day are not driven by the same factors, as the genetic influence on the cortisol awakening response is distinct from the heritability of daytime cortisol levels (Wust et al. 2000). In summary, baseline and stress measurements seem to reflect different processes and cannot be lumped together. To be able to interpret them appropriately in the process of somatization, it is essential that their different meanings will be further elucidated.

Do alterations in HPA axis activity precede somatization?

The question whether alterations in HPA axis activity predict somatization has been subject of a number of prospective studies. First, several studies have examined whether cortisol levels predict FSS or FSD on the long term. Studies focusing on fatigue in specific groups at risk found no predictive role of cortisol in the development of post-infectious unexplained fatigue in 71 primary care patients after three or six months (Candy et al. 2003) or in post-operative unexplained fatigue in 161 patients undergoing elective surgery after two days, three weeks, or six months (Rubin et al. 2005). From a study on fatigue in the general population, however, another picture emerges. Among a group of 4299 former or current civil servants, low cortisol at awakening predicted new-onset fatigue at follow-up approximately 30 months later. The association with new-onset fatigue was independent of factors such as age, gender, waist circumference, smoking, alcohol use, medication use, depressive symptoms, time of waking, sleep duration, sleep disturbances, and stress on the day of sample collection. Furthermore, persistent fatigue assessed two years prior to cortisol measurement was not associated with cortisol secretion, suggesting that the direction of the association is dominantly from lower cortisol leading to fatigue (Kumari et al. 2009). Long-term longitudinal studies have not been restricted to unexplained fatigue, but have also been performed in other FSS or subgroups of FSD. Among a group of subjects psychosocially at high risk for chronic widespread pain, lower morning and higher evening salivary cortisol levels indicating a blunted diurnal rhythm, predicted new onset of chronic widespread pain, a condition related to FM (McBeth et al. 2007). Twenty-four hour UFC (24-h UFC) excretion did not predict development of new-onset FSS in a two year follow-up period in our

previously mentioned population-based study (Chapter 7). It should be noted that fatigue has not been measured in this study. However, also when specifically looking at different bodily clusters bearing resemblance to symptoms of CFS (general FSS cluster), FM (musculoskeletal FSS cluster), and IBS (gastrointestinal FSS cluster), no differential associations with cortisol became apparent.

Other studies have closely examined whether cortisol levels are directly related to the experience of FSS. From these studies, a fairly consistent pattern emerges, in which a flattened diurnal rhythm of cortisol (i.e., lower morning cortisol and higher evening cortisol levels) seems related to somatization. In a sample of 121 participants middle-aged adults, it was found that morning cortisol levels were not predicted by prior-day levels of fatigue and physical symptoms, whereas low morning cortisol predict higher levels of fatigue and physical symptoms later that day. Authors conclude that these results are consistent with a role of cortisol regarding physiological activation and an influence on physical well-being (Adam et al. 2006). In addition, associations between cortisol levels at awakening and one hour after awakening and current pain symptoms in FM have been found, but not at three time points later in the day (McLean et al. 2005).

Can HPA axis activity change or be changed?

Guidelines about variables that should be considered as covariates influencing HPA axis activity have become available for salivary cortisol measurements in epidemiological research, among which are age, body mass index, and smoking (Adam & Kumari 2009b, Vreeburg et al. 2009b). Cortisol levels have also been associated with different somatic and psychiatric conditions (Licinio et al. 1996, Heim et al. 2000a, Lindqvist et al. 2008, Mantella et al. 2008, Vreeburg et al. 2009a). Whereas a meta-analysis points to an association between chronic psychosocial stress and altered HPA axis activity (Miller et al. 2007), we did not find an association between self-reported adverse life events and 24-h UFC using data measured in our population-based cohort (Chapter 9). In a review, several medications with the capacity to influence cortisol levels via different pathways are documented to minimize this potential source of error variance (Granger et al. 2009). Not only corticosteroids, but the possible effects of other hormonal medication such as contraceptives and estrogen replacement therapy, or antidepressants should also be taken into account (Holsboer & Barden 1996, Ambrogio et al. 2008, Roberts et al. 2009b). Alternatively, those medications can also be used to change cortisol levels intentionally as an intervention. Given the widespread actions of this stress responsive system, the high diversity of factors that influence HPA axis activity is not surprising but may generate controversy about its relevance. Like in ANS research, the same conceptual problems apply when deciding which covariates are confounders, mediators, or moderators. For example, we have made adjustments for depressive disorder, a strategy also adopted by other large epidemiological studies (Kumari et al. 2009). However, it is unlikely that depression and somatization constitute completely unrelated

entities (Lowe et al. 2008). Statistical adjustment for such variables may result in a residual independent variable with unknown construct validity. With regard to the pathophysiological meaning of alterations in cortisol levels, available evidence suggest that morning levels are related with experience of FSS later in the day (Adam et al. 2006, Kumari et al. 2009, Fabian et al. 2009), a reason why a lower cortisol awakening response and a flattened diurnal rhythm may be specifically associated with somatization. The exact interpretation of each of the elements of HPA axis activity is subject to debate, but recent reviews are beginning to elucidate the meaning and relevance of the different aspects of the diurnal rhythm (Clow et al. 2004, Chida & Steptoe 2009). The finding of the importance of morning cortisol in somatization may not be surprising, because timely cortisol secretion for mobilization of energy resources is necessary to meet upcoming demands of the day (Buijs et al. 2003). Again, underlying mechanisms may differ by gender or age (i.e., gender or age may be a moderator). For example, HPA axis alterations in FSD are mainly present in females (Chapter 4). Although these findings require replication in new studies, they also urge the need for pathophysiological explanations. In a first genome-wide study on associations with morning serum cortisol, for example, genetic underpinnings of cortisol secretion varied for men and women (Kurina et al. 2005). Regulation of the HPA axis by cortisol is mediated in a complementary fashion by mineralocorticoid- and glucocorticoid receptors (Derijk & de Kloet 2008). Polymorphisms in the mineralocorticoid receptor gene influence the cortisol awakening response after dexamethasone suppression in a sex specific manner (Van Leeuwen et al. 2009). To come closer to a full understanding of the existence of subgroups, the epidemiologic approach to stress responsive system dysfunction and somatization in this thesis should be beneficially followed by basic research into the underlying genetic and molecular mechanisms.

Does manipulation of HPA axis activity change somatization?
It has been investigated whether cortisol level can be changed by treatment or has prognostic value in predicting recovery. Randomized controlled trials have shown that low-dose cortisol replacement therapy lead to short-term reductions in fatigue (McKenzie et al. 1998, Cleare et al. 1999). Although pharmacologically raising levels of cortisol can temporarily alleviate symptoms, it is not recommended as treatment of choice in CFS. Reasons for caution are a rapid loss of efficacy upon discontinuation, the observation that only a minority of patients gain benefit, and that no pre-treatment factors that predicted response to hydrocortisone are identified (Cleare 2004a). Whether the same applies to FM and IBS is unknown. In FM, only one small study in 20 patients showed that 10 mg of prednisone daily was not effective (Clark et al. 1985). In IBS, treatment with corticosteroids has never been tested in a trial (Henningsen et al. 2007).

The effect of psychological and behavioral interventions on the HPA axis has also been tested in FSD. One study in CFS observed that cognitive behavioral therapy

resulted in a significant rise in cortisol levels (Roberts et al. 2009b). Another study found that three weeks of exercise and cognitive behavioral therapy improved the flattened diurnal rhythm of cortisol in FM subjects (Bonifazi et al. 2006). It should be noted that neither study included a control group. Moreover, although these two studies indicated that hypocortisolism in FSD is reversible by treatment, they did specifically examine whether these alterations were due to reducing adverse behavioral consequences, such as sleep disturbances, physical inactivity, or pain experience. A third study demonstrated that lower daily cortisol output and a flattened diurnal rhythm predict a poor response to cognitive behavioral therapy in CFS (Roberts et al. 2009a). Interestingly, it has also been reported that CFS patients who respond less well to cognitive behavioral therapy are the ones who are more persistently physically inactive (Prins et al. 2001). However, the question whether non-responders with hypocortisolism in this study might represent a physically inactive group has not been assessed and remains to be answered. Preliminary evidence for physical inactivity as a cause of hypocortisolism rather than a confounder comes from a small study performed in 18 regularly exercising healthy adults. Subjects were asked to discontinue their regular aerobics lessons for one week. The subset of healthy subjects that developed symptoms of pain and fatigue after exercise deprivation was characterized by lower cortisol levels at baseline (Glass et al. 2004). Authors of this study speculated that the subset of healthy individuals with lower cortisol levels unknowingly exercise regularly to augment the function of HPA axis and thus suppresses symptoms.

Conclusions on the role of HPA axis dysfunction in somatization

The HPA axis is the most extensively investigated stress responsive system in somatization. Hypocortisolism is only associated with CFS and females with FM, whereas the association with IBS seems to be different, notably, a non-statistical significant association in the direction of hypercortisolism. Although HPA axis alterations are not always found, when differences are found, lower morning cortisol levels in combination with a flatted diurnal rhythm seem consistent HPA axis alterations in functional fatigue and musculoskeletal pain. Longitudinal studies suggest that hypocortisolism is a specific risk factor for fatigue and musculoskeletal FSS. Whether hypocortisolism also has a perpetual role in symptom maintenance has not been specifically studied. Based on current available evidence, hypocortisolism is at least a risk factor of subgroups in FSD. The question whether it also is a causal risk factor remains to be answered.

IMMUNE SYSTEM ACTIVITY AND SOMATIZATION

Are alterations in immune system activity correlated with somatization?

As mentioned in the introduction, numerous abnormalities in immune cell quantity and function have been identified in the main three FSD. In a critical review on CFS, no clear differences between patients and healthy controls were apparent in cytokine levels in the studies with the highest quality rates. Other abnormalities could rarely be replicated in more than one paper, making authors of this review conclude that "CFS literature now contains papers with results to support virtually any conclusion about the nature of the immunological abnormalities" (Lyall et al. 2003). Similar heterogeneous immune findings have been reported by reviews on FM (Wallace 2006, Gur & Oktayoglu 2008) and studies on IBS (Liebregts et al. 2007, Kindt et al. 2009). The non-replication of specific immune alterations may be primarily explained by the variety of confounders in these kinds of clinical populations, the co-morbidity with depressive disorder, and the lack of an epidemiologically comparable healthy control group (Pariante 2009). Other markers, such as high-sensitive C-reactive protein (hs-CRP), might be more accurate for studying immune system activation. Indeed, the few available findings addressing hs-CRP in FSD appear more consistent. Patients with CFS have higher levels of hs-CRP compared to healthy controls (Buchwald et al. 1997, Spence et al. 2008, Raison et al. 2009). The latter study also found that a subsyndromal group of fatigued persons had a similar inflammatory profile of those with CFS. Remarkably, adjusted models indicated that subsyndromal fatigue, but not CFS, remained independently associated with hs-CRP after adjustment for factors such as age, gender, body mass index, depressive symptoms, and medication use. Authors were puzzled by this finding and failed to find a satisfactory explanation. Possibly, immune activation in subjects with full-blown CFS has partly another etiology (and is removed by controlling for several covariates) than immune activation in subjects with subsyndromal fatigue. IBS patient had higher CRP levels than healthy controls, but this difference was not statistically significant (Schoepfer et al. 2008). Furthermore, hs-CRP is related to low back pain in a study in young female adults (Shiri et al. 2008) and self-reported pain not due to chronic disease in a study in older adults (Graham et al. 2006). In our population-based study, in which we adjusted for a large range of potential confounding variables such as body mass index, presence of depressive disorder, and physical activity level, no association between hs-CRP and the total number of FSS is observed (Chapter 8). When bodily clusters of FSS were differentially explored, however, hs-CRP was associated with general FSS and pure musculoskeletal FSS, but not with gastrointestinal FSS. When considering hs-CRP as a correlate of somatization, it is important to shine light on the complex relationship between the ANS, HPA axis, and immune system. Immune system activation may be a biomarker for sickness behavior, but also for activity of the ANS and HPA axis. In response to

stress, CRH production in the paraventricular nucleus of the hypothalamus is increased. Besides activation of the HPA axis, CRH acts within the brain to activate the sympathetic nervous system and production of inflammatory cytokines. In turn, high concentrations of cortisol suppress synthesis of pro-inflammatory cytokines, while low cortisol concentrations have been documented to lead to increased release of inflammatory cytokines (Maier & Watkins 1998). In other words, hypocortisolism associated with somatization may result in increased levels of pro-inflammatory cytokines. Both dysregulation of the ANS (Sajadieh et al. 2006, Araujo et al. 2006, Kon et al. 2006, Madsen et al. 2007, Marsland et al. 2007, Singh et al. 2009) and HPA axis (McEwen et al. 1997, Turnbull & Rivier 1999) are proven to be associated with higher concentrations of pro-inflammatory cytokines and CRP, supporting a link between those three stress responsive systems. Although it is well-established that stress responsive systems do not function independently, little is known about specific conditions under which they either reinforce or compensate each other (Turnbull & Rivier 1999). Their joint action in somatization has not been investigated yet.

Do alterations in immune system activity precede somatization?
In the aforementioned population-based study on hs-CRP and FSS, no longitudinal association between hs-CRP and total number of FSS was found (Chapter 8). The differential positive associations between hs-CRP and FSS in the general and pure musculoskeletal bodily clusters present in the cross-sectional analysis, disappeared in the longitudinal analysis for the general bodily cluster but remained for the pure musculoskeletal bodily cluster. We are not aware of other studies investigating whether elevated hs-CRP, or other alterations in immune system activity, precede somatization.

Can immune system activity change or be changed?
A meta-analysis provides only modest evidence of successful immune modulation using psychological interventions; however, authors warn that several methodological issues need to be resolved before any definitive conclusions can be reached (Miller & Cohen 2001). Nonetheless, it is obvious that immune system activity is not static. The list of associations of hs-CRP with conditions that are not apparently inflammatory is virtually endless, and include polymorphisms in glucocorticoid receptor, female gender, low socio-economic status, low fiber consumption, physical inactivity, obesity, hypertension, cataract, and depressive disorder (Kushner et al. 2006). Recently, it has been concluded that a lack of standardization of inclusion and exclusion criteria and assessment of control variables hinders interpretation of studies investigating linkages between psychosocial and behavioral factors and inflammation (O'Connor et al. 2009). Authors provide recommendations as to whether each factor, such as age, hormonal status, socioeconomic status, body mass index, physical inactivity, diet, caffeine, smoking, alcohol, sleep disruption, antidepressants, cardiovascular- and

immunosuppressive medication should be used as an exclusion criterion or controlled for. As mentioned previously, not adequately adjusting for covariates that are confounders may lead to spurious associations, but adjusting for covariates that are mediators may remove true associations. For example, IL-6 and other pro-inflammatory cytokines are not only released by the liver, but are also by adipose tissue (Mohamed-Ali et al. 1997, Pou et al. 2007). In this perspective, body mass index may not be a confounder in the potential relation between hs-CRP and FSS, but may rather be a causal factor inducing low-grade inflammation, sickness behavior, and FSS.

Does manipulation of immune system activity change somatization?
Only one small study has investigated whether treatment of FSD also alters immune system activity as measured by peripheral circulating cytokine levels (Wang et al. 2008). Levels of pro-inflammatory cytokines TNF-α and IL-8 were significantly higher in a group of 20 FM patients compared to healthy controls. After ten days of multidisciplinary pain treatment, elevated TNF-α levels in FM patients had normalized, but IL-8 levels remained significantly higher. Although TNF-α levels were elevated on admission and had normalized after six months of multidisciplinary pain therapy paralleling the decline in pain intensity, there was no direct correlation between TNF-α and pain intensity. Authors conclude that cytokines do not seem to directly cause pain in FM.

Conclusions on the role of immune system dysfunction in somatization
Immune system dysfunction does not seem to be a shared risk factor for FSS and FSD. A cross-sectional association is mainly apparent in persons with functional fatigue and widespread muscle pain, but not in subjects with abdominal complaints as key presenting symptom. The differential associations may be explained by the fact that key FSS in those FSD are related to sickness behavior, mainly to the hyperalgesia and weakness components (Dantzer 2001). This idea is supported by studies on CFS which consistently show that hs-CRP levels are higher in patients compared to healthy controls. Remarkably, the pattern of FSD-specific association between immune system activity and fatigue and musculoskeletal pain seems comparable to the FSD-specific associations of HPA axis dysfunction. Future studies on immune function should take the same variables into account as important in HPA axis studies (i.e., depressive disorder, female gender, medication use, physical inactivity), and preferably study those two systems together. It is uncertain how to measure sickness behavior and clear evidence that hs-CRP is a valid indicator of somatization is lacking (Chapter 8). As it stands now, available results do not rule out the possibility that immune system alterations are related to somatization. It is, however, difficult to take a position with regard to the pathway in which the immune system is involved in somatization, because the optimal measure is unknown, and the direct relation to symptom experience is difficult to assess.

FUTURE DIRECTIONS FOR TESTING DYSFUNCTION OF STRESS RESPONSIVE DYSFUNCTION IN SOMATIZATION: THE NEED TO IDENTIFY SUBGROUPS

By most probable means, somatization can be classified as a multifactorially caused condition. We hypothesized that lower cardiac vagal activity, hypocortisolism, and elevated hs-CRP could be biological risk factors associated with FSS and FSD. In this thesis, we have found some support for those hypotheses. First, lower cardiac vagal activity was associated with FSS in general and FSD. Second, hypocortisolism was associated with FM in females and with CFS. Third, elevated hs-CRP was associated with musculoskeletal FSS. However, dysfunction of stress responsive systems is not a risk factor that applies to somatization in general. Specific alterations in stress responsive systems may play a role in the etiology of some types of somatization, but several complexities appear.

First, a specific alteration in stress responsive system function might be a risk factor for more than one cluster of FSS or FSD (e.g., hypocortisolism might be involved in FM and CFS or hs-CRP elevations may be associated with musculoskeletal and fatigue-related FSS). Based on our research on HPA axis and immune system activity, we conclude that CFS and FM seem more similar than CFS and IBS or FM and IBS. This finding confirms previous latent class analyses indicating that CFS and FM are not particularly distinctive but are characterized by greater similarities than differences (Sullivan et al. 2002c). From a causal point of view, this may imply that these symptom clusters or FSD share their etiology. The key presenting fatigue and musculoskeletal pain symptoms may be provoked by alterations of these stress responsive systems.

Second, opposing alterations in the same stress responsive system may be related to different FSD (e.g., hypocortisolism is associated with CFS, whereas hypercortisolism may be associated with IBS). From a causal point of view, this may imply that different patterns of dysfunction in the same stress responsive system lead to differences in symptom experience. Pathophysiological mechanisms that underlie this possibility remain to be established.

Third, alterations in stress responsive system function may only be relevant in subgroups of patients within FSD (e.g., hypocortisolism seems only relevant in female but not male FM patients and hypocortisolism is only relevant in CFS patients with a history of childhood trauma as opposed to those without such a history (Heim et al. 2009)). It should be noted that comparable stress responsive system alterations may be causal in one subgroup but consequential in another subgroup. Moreover, stress responsive system dysfunction on its turn is also influenced by several factors, complicating the elucidation of clear-cut pathways.

Fourth, opposing alterations in stress responsive system function may be found within the same FSS or FSD (e.g., lower cardiac vagal activity is related to FSS in young adults, whereas higher cardiac vagal activity is related to FSS in middle-aged to old adults). From a causal point of view, this may imply that different pathophysiological processes in different groups can lead to the same outcome in symptom experience. Again, pathophysiological mechanisms that underlie this possibility remain to be established.

To successfully proceed in this complex field, future research faces the challenge of devising studies that include a theoretical perspective as to why alterations in stress responsive system function might exert influence on somatization. In addition, carefully choosing the optimal methodological strategy is essential. We have the following recommendations for future research on dysfunction of stress responsive systems in the etiology of somatization.

Research design

The same etiological factor is often only researched in all FSD separately. Researchers should not only focus on the diagnostic criteria of separate FSD, but also routinely measure somatization more broadly with a symptom checklist (i.e., multi-symptom assessment). Moreover, they should aim to include patients with different types of FSD in one study design in order to systematically appraise differences and similarities (i.e., multi-disorder assessment). When choosing to use case-control design, effort should be made to select controls from the same population as the cases. Selection bias can be diminished by the use of nested case–control studies in population-based cohorts. Furthermore, power calculations should be made before conducting a study. To date, no case-control study has been published that is adequately powered to find effect sizes that are usual in this field. For example, the effect size of lower cardiac vagal activity in FSD compared to healthy controls was $d = 0.32$. When performing a case-control study, a sample size of 155 in each group will have 80% power to detect an effect size of this magnitude using a two group *t*-test with a 0.05 two-sided significance level. To detect an effect of this magnitude with 90% power, a sample size of 207 in each group is required. The prevalence of the main FSD in the population roughly ranges from 1 - 5% (Kato et al. 2009). When performing a population-based nested case-control study, consequently, at least a cohort of roughly 3000 - 15000 participants is required. Regarding studies on HPA axis activity, it has been calculated that even for relatively prevalent conditions, many existing cortisol studies would not be sufficiently powered to detect a significance difference at the 0.05 two-sided significance level (Adam & Kumari 2009). This calculation also applies to the topic of HPA axis activity and somatization. For example, the effect size of hypocortisolism in CFS subjects compared to healthy controls was $d = 0.14$. A sample size of 802 in each group will have 80% power to detect an effect size of this magnitude using a two group *t*-test with a 0.05 two-sided significance level. A sample size of 1074 subjects in each group is necessary to have sufficient

power to detect this effect with 90% power. The estimated prevalence of CFS ranges from 0.1 - 1% (Prins et al. 2006, Harvey et al. 2008). When performing a population-based nested case-control study, a cohort of roughly 8000 - 80000 participants is needed. Thus, larger studies are required for a decisive answer on the role of HPA axis dysfunction in FSD.

Two directions of research seem most promising: research on population-based cohorts and research on intervention studies. Population-based studies can use principles of life course epidemiology, that is studying long-term effects on later somatization risk of biological, psychological, and social exposures influencing stress responsive systems during gestation, childhood, adolescence, and adult life (Kuh et al. 2003). Although at this point in time it may be premature to propose detailed models of stress responsive system dysfunction in somatization, this thesis has provided some directions that seem particularly interesting. Adopting a life course approach presents major challenges for both the design and analysis, as access to a birth cohort study with repeated measurements of all these variables may be unrealistic in the coming years. Furthermore, it is costly to study stress responsive system function in sufficiently large samples.

Meanwhile, another interesting way to promote understanding of dysfunction of stress responsive systems as a causal risk factor of the etiology of somatization is doing research on treatment studies. Using evidence-based treatments for FSS or FSD that are currently available (Henningsen et al. 2007), this strategy may provide information as to whether measures of stress responsive system function predict treatment response or change when somatization improves. For example, hypocortisolism and a flattened diurnal release of cortisol were associated with a poorer response to cognitive behavioral therapy in CFS (Roberts et al. 2009a). However, it is important to realize that predictors of remission might differ from predictors of disease onset. Nevertheless, patients' neuroendocrine profile may be relevant in choosing the optimal treatment strategy. Research on treatment studies may contribute to decide correctly for whom, when, and how to intervene to prevent the onset of somatization or to facilitate recovery (Kraemer et al. 2001).

Ultimately, a risk factor can only be proven to be causal when it has the ability to change the outcome in randomized controlled trials. As a second step, the benefits of treatments targeting stress responsive system dysfunction could be studied. Importantly, ignoring strong moderators of treatment response may lead to inclusion of many subjects for whom the interventions are not appropriate, or perhaps are even harmful, and may attenuate effect sizes. In fact, weak effects associated with various treatments for FSD (Sumathipala 2007, Henningsen et al. 2007) may be due to lack of information on moderators and mediators of treatment. Instead, interventions should be targeted to relevant subgroups (Kraemer et al. 2001). Taking a broader perspective, a classification of

somatization based on risk factors appears more logical than a classification mainly based on key presenting or number of symptoms.

Measurement of stress responsive systems, somatization, and covariates

Measurement of somatization is a controversial issue, because subjective symptom experience has to be assessed. Eight problems in sampling and measurement of symptoms have been recognized (Kroenke 2001). First, the exceptionally high prevalence of many symptoms is essential to consider when selecting appropriate cases and controls. Most persons who have FSS never report them to a physician, resulting in a large reservoir of FSS in the community (Kroenke & Price 1993, Ihlebaek et al. 2002). There does not appear to be a clear-cut FSS count threshold that justifies a cut-point; however, operational cut-points are also established for other continuous psychiatric disorders (e.g., depressive disorder, anxiety disorder) and medical disorders (e.g., hypertension, diabetes mellitus). Nevertheless, besides the number of FSS, additional criteria on psychological, biological, and social factors may be required for the diagnosis of somatization (Kroenke et al. 2007). Second, the type of population studied, such as persons from the community, consulters from a primary care clinic, or patients from a tertiary centre can greatly influence the severity and other characteristics of the symptoms being evaluated. Although we found that a cut-off of four FSS in the previous year distinguishes clinically relevant somatization from a healthy condition in our population cohort (Chapter 5), this number might be different in other populations. Third, symptom detection relies on three types of assessment, namely, chart review, symptom checklists, or spontaneous report. Each method identifies FSS of a different nature and threshold that should be carefully considered. We have measured the number of FSS with the somatization section of the Composite International Diagnostic Interview (CIDI) (Andrews & Peters 1998). Other structured interviews, such as the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (Wing et al. 1990), or symptom checklists, such as the somatic items of the 90-items Symptom Checklist (SCL-90) (Derogatis et al. 1973) and the Screening for Somatoform Symptoms (SOMS) (Rief et al. 1997), may require a different number of FSS to identify clinically relevant somatization. Fourth, temporal factors, such as recency of onset, remittance, and duration should be assessed. Counting lifetime symptoms appear especially problematic. Besides it impracticability in clinical settings, lifetime recall of specific symptoms has a poor reliability between two time-points (Leiknes et al. 2006). Moreover, the induction time between exposure to dysfunction of stress responsive systems and initiation of somatization is largely unexplored, while the strength of an association between exposure and disease can be diluted or missed if measured in the wrong time frame (Kuh et al. 2003). Fifth, severity of FSS should be assessed with symptom-specific scales as well as evaluation of general functional status. For example, working status or functional disability could be measured. Sixth, recognition of a symptom's cause can be strengthened by explicit criteria

and uniform assessment protocols. The use of the CIDI may have resulted in an underestimation of the true amount of FSS in our population, since patients presumably tend to interpret FSS as medically explained rather than the opposite (Robins et al. 1982). FSS assessment with multiple rating system or multiple raters would help to gain knowledge on the clinical usefulness of each method. On the other side, the need of excluding somatic diseases can be questioned. We conducted sensitivity analyses after exclusion of participants with somatic disease, such as cardiovascular disease, rheumatoid arthritis, chronic obstructive pulmonary disease, inflammatory bowel disorders, malignancy, and the results remained essentially the same (Chapters 6 - 8). This may be due to the phenomenon that the extent to which people suffer from medically explained symptoms and disorders is also largely explained by processes underlying somatization (Kroenke & Rosmalen 2006). Moreover, associations with higher body mass index, abstinence of alcohol, physical inactivity, lower socio-economic status, and psychological distress are comparable for FSS and medically explained symptoms (Kingma et al. 2009). Seventh, co-morbidity that should be assessed includes medical and psychiatric disorders. For example, the presence of a primary underlying psychiatric disorder could lead to an amplification of or excessive worry about bodily sensations. Furthermore, whilst many researchers agree FSD have at least a common core, most original research towards these disorders focus on one single disorder, and the same etiological factor is often only researched in FSD separately. Especially in case of FSD, symptoms of other FSS and FSD should be examined to gain knowledge on differential associations specific bodily clusters with of stress responsive system dysfunction. Eighth, since hard end-points are uncommon, alternative outcomes that might be measured include symptom alleviation or health care costs (Kroenke 2001). There is no gold standard in the field of somatization; even among experts trying to reach consensus divergent views are usually held (Kroenke et al. 2007). Taking those eight points into account, however, an operational or working definition used to facilitate more uniform classification and reproducibility in subsequent studies may be the bronze standard. The importance of uniform assessment of stress responsive systems and careful measurement of covariates that influence the association between dysfunction of stress responsive systems and somatization has been discussed previously.

Data-analysis

Future studies could employ methods that appropriately account for the time-varying nature of the association by repeatedly assessing activity of stress responsive systems, somatization, and covariates like body mass index, depression, anxiety, and health behaviors, such as smoking, alcohol use, and physical inactivity, and factors like childhood trauma and other psychosocial stressors (Christenfeld et al. 2004). With regard to confounders, mediators, and moderators, it is difficult to decide how to build the models, partly because of a

lack of high quality data from sufficiently large studies, and partly because of the difficulty in identifying underlying biologic phenomena from statistical models. There is no substitute for carefully considering and explicitly stating the rationale of included covariates (Babyak 2004). Post-hoc or sensitivity analyses can contribute to knowledge of the effects of different decisions.

Previous studies have generally applied *t*-test comparisons, analysis of variance, and regression models. Those statistical models that make parametric assumptions or primarily model linear changes and may limit better understanding of this multifactorial, complex field. Statistical models that lend themselves equally well to increases as decreases in outcomes and do not make assumptions about linearity would therefore be a fundamental advantage. Ideally, future research would use multi-level models for change that focus on within individual changes (i.e., how does each person change over time, thus, individual growth trajectories) and interindividual differences in change (i.e., what predict differences across people in their changes). Multi-level models for change needs three or more waves of data, but allows time-unstructured and unbalanced number of measurements waves (Snijders & Bosker 1999, Singer & Willett 2003). This statistical technique can not only be applied to trajectories towards disease, but also to trajectories towards health during treatment studies.

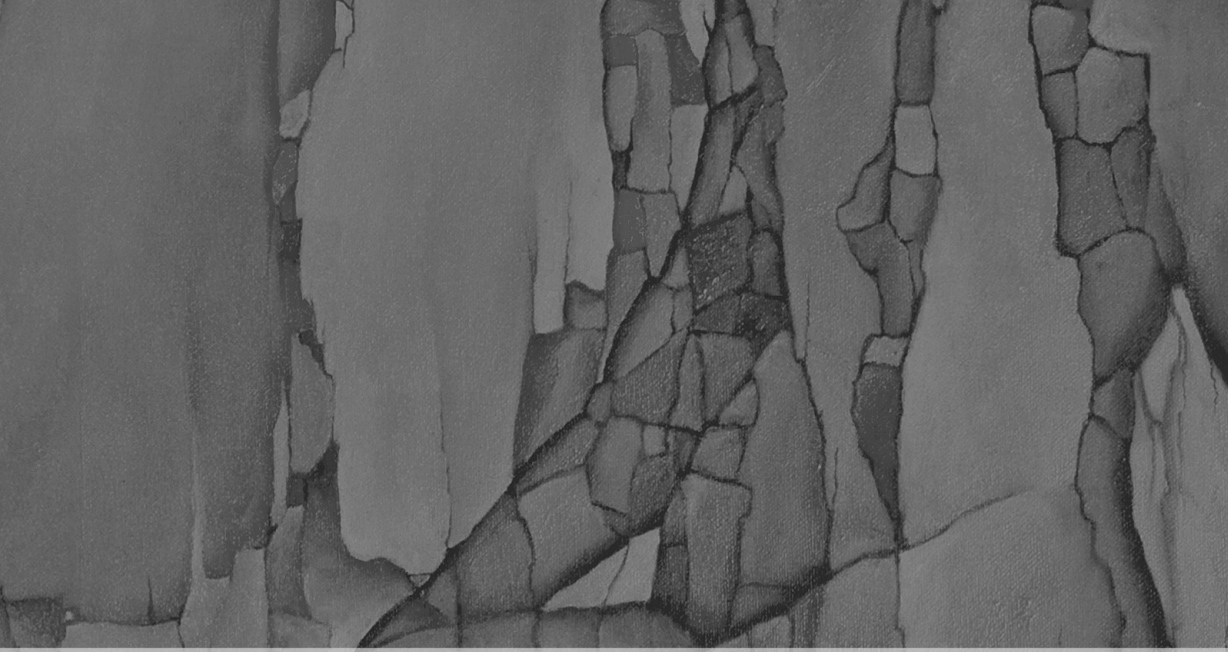
Interpretations and clinical relevance

Statistical significance does not inform us about the potency of a risk factor (Kraemer et al. 1997), and a relationship may be even uninteresting if its effect size is small. Thus, when stress responsive system is found to be a causal risk factor for somatization, its value in clinical or policy applications remains to be evaluated. According to Cohen's conventions, the magnitude of significant associations between stress responsive system function and FSS is usually small (Cohen 1988). However, the terms small, medium, and large are relative, not only to each other, but even more particularly to the area of research. With regard to somatization, several explaining mechanisms, of which some seem intuitively more attractive than dysfunction of stress responsive systems, have been proposed. Examples of such mechanisms are somatosensory amplification, dysfunctional health cognitions, symptom-led activity patterns, and ineffective coping (Rief & Broadbent 2007a, Deary et al. 2007). As is the case in stress responsive system research, associations in those research areas are usually small and not unambiguous. Thus, etiological studies focused on psychological and social factors also support the idea that disturbances in generation, attention, interpretation, and behavioral processes may all contribute to somatization, but in varying extents in different persons. Research should focus on identifying those disease trajectories.

Furthermore, rather than relying on Cohen's conventions, effect sizes found in this thesis may be better compared with widely accepted clinically important effect sizes of associations in a related research domain. An example from a related research domain is that of psychiatric co-morbidity and FSD. The effect size of the association between hypocortisolism and CFS in our meta-analysis was d = 0.14 (Chapter 3). In comparison, an effect size of d = 0.33 was found in a meta-analysis on the association between current depressive disorder and FSD (Henningsen et al. 2003). Although the latter effect size is twice as large, depression and somatization were both assessed by self-report, potentially inflating the strength of the association by shared method variance (Tepper & Tepper 1993), whereas measurement of cortisol does not have shared method variance with somatization. Furthermore, a large population-based study on cortisol in depressive disorder found a comparable effect size. Patients with depressive disorder had significantly higher cortisol awakening response compared with control subjects (d = 0.15 - 0.25) (Vreeburg et al. 2009a). Potential causes of small effect sizes are heterogeneity in etiological mechanisms, which are inherent to associations typically studied in psychosomatic research. The fact that an effect size is an average can be important. Whereas a small overall effect size may raise questions about the clinical relevance of the association, this problem may disappear when studying relevant subgroups. The importance of finding subgroups is illustrated by the meta-analysis on HPA axis activity in FSD, in which the effects size of hypocortisolism was substantially larger in females with FM, notably, d = 0.26, compared to the effect size of hypocortisolism in all FM patients (d = 0.10).

CONCLUDING REMARKS

In conclusion, although not a *condition sine qua non* for somatization in general, stress responsive system dysfunction may be pivotal in the etiology and treatment strategy in subgroups of patients. Such subgroups need to be better identified. As several studies begin to gather multiple waves of stress responsive system data and important covariates over the course of many years, the role of and time scales over which changes contribute to somatization should be increasingly illuminated. This approach likely represents the best strategy by which we can improve our understanding of this association, and ultimately, achieve benefit for patients.



REFERENCES

Aaron LA, Buchwald D. 2001. A review of the evidence for overlap among unexplained clinical conditions. *Ann. Intern. Med.* 134(9 Pt 2):868-81
Aaron LA, Burke MM, Buchwald D. 2000. Overlapping conditions among patients with chronic fatigue syndrome, fibromyalgia, and temporomandibular disorder. *Arch. Intern. Med.* 160(2):221-7
Adam EK, Hawkley LC, Kudielka BM, Cacioppo JT. 2006. Day-to-day dynamics of experience--cortisol associations in a population-based sample of older adults. *Proc. Natl. Acad. Sci. U. S. A.* 103(45):17058-63
Adam EK, Kumari M. 2009. Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology.* 34(10):1423-36
Adeyemi EO, Desai KD, Towsey M, Ghista D. 1999. Characterization of autonomic dysfunction in patients with irritable bowel syndrome by means of heart rate variability studies. *Am. J. Gastroenterol.* 94(3):816-23
Adler GK, Kinsley BT, Hurwitz S, Mossey CJ, Goldenberg DL. 1999. Reduced hypothalamic-pituitary and sympathoadrenal responses to hypoglycemia in women with fibromyalgia syndrome. *Am. J. Med.* 106(5):534-43
Agelink MW, Majewski TB, Andrich J, Mueck-Weymann M. 2002. Short-term effects of intravenous benzodiazepines on autonomic neurocardiac regulation in humans: a comparison between midazolam, diazepam, and lorazepam. *Crit. Care Med.* 30(5):997-1006
Aggarwal VR, McBeth J, Zakrzewska JM, Lunt M, Macfarlane GJ. 2006. The epidemiology of chronic syndromes that are frequently unexplained: do they have common associated factors? *Int. J. Epidemiol.* 35(2):468-76
Albert MA, Glynn RJ, Ridker PM. 2003. Alcohol consumption and plasma concentration of C-reactive protein. *Circulation.* 107(3):443-7
Altemus M, Dale JK, Michelson D, Demitrack MA, Gold PW, Straus SE. 2001. Abnormalities in response to vasopressin infusion in chronic fatigue syndrome. *Psychoneuroendocrinology.* 26(2):175-88
Altman DG. 1996. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ.* 313(7057):570-1
Altman DG. 2000. Statistics in medical journals: some recent trends. *Stat. Med.* 19(23):3275-89
Altman DG, Lyman GH. 1998. Methodological challenges in the evaluation of prognostic factors in breast cancer. *Breast Cancer Res. Treat.* 52(1-3):289-303
Ambrogio AG, Pecori GF, Cavagnini F. 2008. Drugs and HPA axis. *Pituitary.* 11(2):219-29
American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. 1994.
Andrews G, Peters L. 1998. The psychometric properties of the Composite International Diagnostic Interview. *Soc. Psychiatry Psychiatr. Epidemiol.* 33(2):80-8
Antonijevic IA. 2006. Depressive disorders -- is it time to endorse different pathophysiologies? *Psychoneuroendocrinology* 31(1):1-15
Araujo F, Antelmi I, Pereira AC, Latorre MR, Grupi CJ et al. 2006. Lower heart rate variability is associated with higher serum high-sensitivity C-reactive protein concentration in healthy individuals aged 46 years or more. *Int. J. Cardiol.* 107(3):333-7
Ataoglu S, Ozcetin A, Yildiz O, Ataoglu A. 2003. Evaluation of dexamethasone suppression test in fibromyalgia patients with or without depression. *Swiss. Med. Wkly.* 19;133(15-16):241-4
Babyak MA. 2004. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom. Med.* 66(3):411-21
Badrick E, Kirschbaum C, Kumari M. 2007. The relationship between smoking status and cortisol secretion. *J. Clin. Endocrinol. Metab.* 92(3):819-24
Baron RM, Kenny DA. 1986. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51(6):1173-82
Barsky AJ. 2001. Palpitations, arrhythmias, and awareness of cardiac activity. *Ann. Intern. Med.* 134(9 Pt 2):832-7
Barsky AJ, Borus JF. 1999. Functional somatic syndromes. *Ann. Intern. Med.* 130(11):910-21
Barsky AJ, Orav EJ, Bates DW. 2005. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. *Arch. Gen. Psychiatry.* 62(8):903-10
Barth J, Schumacher M, Herrmann-Lingen C. Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. 2004. *Psychosom. Med*. 66(6):802-13

Bax L, Ikeda N, Fukui N, Yaju Y, Tsuruta H, Moons KG. 2009. More than numbers: the power of graphs in meta-analysis. *Am. J. Epidemiol.* 169(2):249-55
Bax L, Yu LM, Ikeda N, Moons KG. 2007. A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC. Med. Res. Methodol.* 7:40.:40
Bazzano LA, He J, Muntner P, Vupputuri S, Whelton PK. 2003. Relationship between cigarette smoking and novel risk factors for cardiovascular disease in the United States. *Ann. Intern. Med.* 138(11):891-7
Bearn J, Allain T, Coskeran P, Munro N, Butler J et al. 1995. Neuroendocrine responses to d-fenfluramine and insulin-induced hypoglycemia in chronic fatigue syndrome. *Biol. Psychiatry* 37(4):245-52
Begg CB, Mazumdar M. 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics.* 50(4):1088-101
Bensing JM, Verhaak PF. 2006. Somatisation: a joint responsibility of doctor and patient. *Lancet.* 367(9509):452-4
Berlin JA. 1995. Invited commentary: benefits of heterogeneity in meta-analysis of data from epidemiologic studies. *Am. J. Epidemiol.* 142(4):383-7
Berntson GG, Bigger JT, Jr., Eckberg DL, Grossman P, Kaufmann PG et al. 1997. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology.* 34(6):623-48
Berrahmoune H, Herbeth B, Lamont JV, Lambert D, Blankenberg S et al. 2007. Association of classical and related inflammatory markers with high-sensitivity C-reactive protein in healthy individuals: results from the Stanislas cohort. *Clin. Chem. Lab Med.* 45(10):1339-46
Bigger JT, Fleiss JL, Rolnitzky LM, Steinman RC. 1993. The ability of several short-term measures of RR variability to predict mortality after myocardial infarction. *Circulation.* 88(3):927-34
Black S, Kushner I, Samols D. 2004. C-reactive Protein. *J. Biol. Chem.* 279(47):48487-90
Blanchard EB, Scharff L. 2002. Psychosocial aspects of assessment and treatment of irritable bowel syndrome in adults and recurrent abdominal pain in children. *J. Consult Clin. Psychol.* 70(3):725-38
Blechert J, Michael T, Grossman P, Lajtman M, Wilhelm FH. 2007. Autonomic and respiratory characteristics of posttraumatic stress disorder and panic disorder. *Psychosom. Med.* 69(9):935-43
Bleil ME, Gianaros PJ, Jennings JR, Flory JD, Manuck SB. 2008. Trait negative affect: toward an integrated model of understanding psychological risk for impairment in cardiac autonomic function. *Psychosom. Med.* 70(3):328-37
Bohmelt AH, Nater UM, Franke S, Hellhammer DH, Ehlert U. 2005. Basal and stimulated hypothalamic-pituitary-adrenal axis activity in patients with functional gastrointestinal disorders and healthy controls. *Psychosom. Med.* 67(2):288-94
Boneva RS, Decker MJ, Maloney EM, Lin JM, Jones JF et al. 2007. Higher heart rate and reduced heart rate variability persist during sleep in chronic fatigue syndrome: a population-based study. *Auton. Neurosci.* 137(1-2):94-101
Bonifazi M, Suman AL, Cambiaggi C, Felici A, Grasso G et al. 2006. Changes in salivary cortisol and corticosteroid receptor-alpha mRNA expression following a 3-week multidisciplinary treatment program in patients with fibromyalgia. *Psychoneuroendocrinology.* 31(9):1076-86
Borenstein M, Hedges LV, Higgins JP, Rothstein HR. 2009. *Introduction to Meta-Analysis,* West Sussex, Wiley.
Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP et al. 2003. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ.* 326(7379):41-4
Bouma EM, Riese H, Ormel J, Verhulst FC, Oldehinkel AJ. 2009. Adolescents' cortisol responses to awakening and social stress; effects of gender, menstrual phase and oral contraceptives. The TRAILS study. *Psychoneuroendocrinology.* 34(6):884-93
Bradburn MJ, Deeks JJ, Altman DG. 1998. METAN -an alternative meta-analysis command. *Stata Tech. Bull.* 44:1-15
Britton A, Shipley M, Malik M, Hnatkova K, Hemingway H, Marmot M. 2007. Changes in heart rate and heart rate variability over time in middle-aged men and women in the general population (from the Whitehall II Cohort Study). *Am. J. Cardiol.* 100(3):524-7
Broderick JE, Arnold D, Kudielka BM, Kirschbaum C. 2004. Salivary cortisol sampling compliance: comparison of patients and healthy volunteers. *Psychoneuroendocrinology.* 29(5):636-50

Brosschot JF, Gerin W, Thayer JF. 2006. The perseverative cognition hypothesis: a review of worry, prolonged stress-related physiological activation, and health. *J. Psychosom. Res.* 60(2):113-24
Brown RJ. 2007. Introduction to the special issue on medically unexplained symptoms: background and future directions. *Clin. Psychol. Rev.* 27(7):769-80
Brugha T, Bebbington P, Tennant C, Hurry J. 1985. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychol. Med.* 15(1):189-94
Brugha TS, Cragg D. 1990. The List of Threatening Experiences: the reliability and validity of a brief life events questionnaire. *Acta Psychiatr. Scand.* 82(1):77-81
Buchwald D, Wener MH, Pearlman T, Kith P. 1997. Markers of inflammation and immune activation in chronic fatigue and chronic fatigue syndrome. *J. Rheumatol.* 24(2):372-6
Buckley TC, Kaloupek DG. 2001. A meta-analytic examination of basal cardiovascular activity in posttraumatic stress disorder. *Psychosom. Med.* 63(4):585-94
Buckley TM, Schatzberg AF. 2005. On the interactions of the hypothalamic-pituitary-adrenal (HPA) axis and sleep: normal HPA axis activity and circadian rhythm, exemplary sleep disorders. *J Clin. Endocrinol. Metab.* 90(5):3106-14
Buffington CA. 2009. Developmental influences on medically unexplained symptoms. *Psychother. Psychosom.* 78(3):139-44
Buijs RM, van Eden CG, Goncharuk VD, Kalsbeek A. 2003. The biological clock tunes the organs of the body: timing by hormones and the autonomic nervous system. *J. Endocrinol.* 177(1):17-26
Burke HM, Davis MC, Otte C, Mohr DC. 2005. Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology.* 30(9):846-56
Burr RL. 2007. Interpretation of normalized spectral heart rate variability indices in sleep research: a critical review. *Sleep.* 30(7):913-9
Burr RL, Jarrett ME, Cain KC, Jun SE, Heitkemper MM. 2009. Catecholamine and cortisol levels during sleep in women with irritable bowel syndrome. *Neurogastroenterol. Motil.* 21(11):1148-55
Burton C. 2003. Beyond somatisation: a review of the understanding and treatment of medically unexplained physical symptoms (MUPS). *Br. J. Gen. Pract.* 53(488):231-9
Burton C, Weller D, Sharpe M. 2009. Functional somatic symptoms and psychological states: an electronic diary study. *Psychosom. Med.* 71(1):77-83
Calis M, Gokce C, Ates F, Ulker S, Izgi HB et al. 2004. Investigation of the hypothalamo-pituitary-adrenal axis (HPA) by 1 microg ACTH test and metyrapone test in patients with primary fibromyalgia syndrome. *J. Endocrinol. Invest.* 27(1):42-6
Call-Hosenfeld JS, Goldenberg DL, Hurwitz S, Adler GK. 2003. Growth hormone and insulin-like growth factor-1 concentrations in women with fibromyalgia. *J. Rheumatol.* 30(4):809-14
Candy B, Chalder T, Cleare AJ, Peakman A, Skowera A et al. 2003. Predictors of fatigue following the onset of infectious mononucleosis. *Psychol. Med.* 33(5):847-55
Cannon JG, Angel JB, Abad LW, O'Grady J, Lundgren N et al. 1998. Hormonal influences on stress-induced neutrophil mobilization in health and chronic fatigue syndrome. *J. Clin. Immunol.* 18(4):291-8
Carrasco S, Gonzalez R, Jimenez J, Roman R, Medina V, Azpiroz J. 1998. Comparison of the heart rate variability parameters obtained from the electrocardiogram and the blood pressure wave. *J. Med. Eng. Technol.* 22(5):195-205
Carter FA, Frampton CM, Mulder RT. 2006. Cesarean section and postpartum depression: a review of the evidence examining the link. *Psychosom. Med.* 68(2):321-30
Carvalho LS, Correa H, Silva GC, Campos FS, Baiao FR et al. 2008. May genetic factors in fibromyalgia help to identify patients with differentially altered frequencies of immune cells? *Clin. Exp. Immunol.* 154(3):346-52
Catley D, Kaell AT, Kirschbaum C, Stone AA. 2000. A naturalistic evaluation of cortisol secretion in persons with fibromyalgia and rheumatoid arthritis. *Arthritis Care Res.* 13(1):51-61
Cesari M, Penninx BW, Newman AB, Kritchevsky SB, Nicklas BJ et al. 2003. Inflammatory markers and onset of cardiovascular events: results from the Health ABC study. *Circulation.* 108(19):2317-22
Cevik R, Gur A, Acar S, Nas K, Sarac AJ. 2004. Hypothalamic-pituitary-gonadal axis hormones and cortisol in both menstrual phases of women with chronic fatigue syndrome and effect of depressive mood on these hormones. *BMC. Musculoskelet. Disord.* 5:47
Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S et al. 1993. Development of a fatigue scale. *J. Psychosom. Res.* 37(2):147-53

Chang L, Sundaresh S, Elliott J, Anton PA, Baldi P et al. 2009. Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis in irritable bowel syndrome. *Neurogastroenterol. Motil.* 21(2):149-59
Chida Y, Hamer M, Steptoe A. 2008a. A bidirectional relationship between psychosocial factors and atopic disorders: a systematic review and meta-analysis. *Psychosom. Med.* 70(1):102-16
Chida Y, Steptoe A. 2008b. Positive psychological well-being and mortality: a quantitative review of prospective observational studies. *Psychosom. Med.* 70(7):741-56
Chida Y, Steptoe A. 2009. Cortisol awakening response and psychosocial factors: a systematic review and meta-analysis. *Biol. Psychol.* 80(3):265-78
Christenfeld NJ, Sloan RP, Carroll D, Greenland S. 2004. Risk factors, confounding, and the illusion of statistical control. *Psychosom. Med.* 66(6):868-75
Chrousos GP, Gold PW. 1992. The concepts of stress and stress system disorders. Overview of physical and behavioral homeostasis. *JAMA.* 267(9):1244-52
Cho HJ, Hotopf M, Wessely S. The placebo response in the treatment of chronic fatigue syndrome: a systematic review and meta-analysis. 2005. *Psychosom. Med.* 67(2):301-13.
Clark S, Tindall E, Bennett RM. 1985. A double blind crossover trial of prednisone versus placebo in the treatment of fibrositis. *J. Rheumatol.* 12(5):980-3
Clauw DJ, Chrousos GP. 1997. Chronic pain and fatigue syndromes: overlapping clinical and neuroendocrine features and potential pathogenic mechanisms. *Neuroimmunomodulation.* 4(3):134-53
Cleare AJ. 2003. The neuroendocrinology of chronic fatigue syndrome. *Endocr. Rev.* 24(2):236-52
Cleare AJ. 2004a. The HPA axis and the genesis of chronic fatigue syndrome. *Trends Endocrinol. Metab.* 15(2):55-9
Cleare AJ, Bearn J, Allain T, McGregor A, Wessely S et al. 1995. Contrasting neuroendocrine responses in depression and chronic fatigue syndrome. *J. Affect. Disord.* 34(4):283-9
Cleare AJ, Blair D, Chambers S, Wessely S. 2001a. Urinary free cortisol in chronic fatigue syndrome. *Am. J. Psychiatry* 158(4):641-3
Cleare AJ, Heap E, Malhi GS, Wessely S, O'Keane V, Miell J. 1999. Low-dose hydrocortisone in chronic fatigue syndrome: a randomised crossover trial. *Lancet.* 353(9151):455-8
Cleare AJ, Miell J, Heap E, Sookdeo S, Young L et al. 2001b. Hypothalamo-pituitary-adrenal axis dysfunction in chronic fatigue syndrome, and the effects of low-dose hydrocortisone therapy. *J. Clin. Endocrinol. Metab* 86(8):3545-54
Cleare AJ, O'Keane V, Miell JP. 2004b. Levels of DHEA and DHEAS and responses to CRH stimulation and hydrocortisone treatment in chronic fatigue syndrome. *Psychoneuroendocrinology.* 29(6):724-32
Clemens JQ. 2008. Male and female pelvic pain disorders--is it all in their heads? *J. Urol.* 179(3):813-4
Clow A, Thorn L, Evans P, Hucklebridge F. 2004. The awakening cortisol response: methodological issues and significance. *Stress.* 7(1):29-37
Cohen H, Benjamin J. 2006. Power spectrum analysis and cardiovascular morbidity in anxiety disorders. *Auton. Neurosci.* 128(1-2):1-8
Cohen H, Loewenthal U, Matar M, Kotler M. 2001a. Association of autonomic dysfunction and clozapine. Heart rate variability and risk for sudden death in patients with schizophrenia on long-term psychotropic medication. *Br. J. Psychiatry.* 179:167-71.:167-71
Cohen H, Neumann L, Alhosshle A, Kotler M, bu-Shakra M, Buskila D. 2001b. Abnormal sympathovagal balance in men with fibromyalgia. *J. Rheumatol.* 28(3):581-9
Cohen H, Neumann L, Shore M, Amir M, Cassuto Y, Buskila D. 2000. Autonomic dysfunction in patients with fibromyalgia: application of power spectral analysis of heart rate variability. *Semin. Arthritis Rheum.* 29(4):217-27
Cohen J. 1988. *Statistical power analysis for the behavioral sciences,* New York, Academic Press.
Colditz GA, Burdick E, Mosteller F. 1995. Heterogeneity in meta-analysis of data from epidemiologic studies: a commentary. *Am. J. Epidemiol.* 142(4):371-82
Collins SM, Karasek RA, Costas K. 2005. Job strain and autonomic indices of cardiovascular disease risk. *Am. J. Ind. Med.* 48(3):182-93
Cordero DL, Sisto SA, Tapp WN, LaManca JJ, Pareja JG, Natelson BH. 1996. Decreased vagal power during treadmill walking in patients with chronic fatigue syndrome. *Clin. Auton. Res.* 6(6):329-33
Cordova AM. 2005. A power spectral analysis of heart rate variability in women with fibromyalgia during induced stress. Dissertation. Alliant International University, San Diego.

Counsell C. 1997. Formulating questions and locating primary studies for inclusion in systematic reviews. *Ann. Intern. Med.* 127(5):380-7
Craig AD. 2002. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3(8):655-66
Creed F, Barsky A. 2004. A systematic review of the epidemiology of somatisation disorder and hypochondriasis. *J. Psychosom. Res.* 56(4):391-408
Crofford LJ, Pillemer SR, Kalogeras KT, Cash JM, Michelson D et al. 1994. Hypothalamic-pituitary-adrenal axis perturbations in patients with fibromyalgia. *Arthritis Rheum.* 37(11):1583-92
Crofford LJ, Young EA, Engleberg NC, Korszun A, Brucksch CB et al. 2004. Basal circadian and pulsatile ACTH and cortisol secretion in patients with fibromyalgia and/or chronic fatigue syndrome. *Brain Behav. Immun.* 18(4):314-25
Dahlgren A, Kecklund G, Theorell T, Akerstedt T. 2009. Day-to-day variation in saliva cortisol-Relation with sleep, stress and self-rated health. *Biol. Psychol.* 82(2):149-55
Danese A, Moffitt TE, Pariante CM, Ambler A, Poulton R, Caspi A. 2008. Elevated inflammation levels in depressed adults with a history of childhood maltreatment. *Arch. Gen. Psychiatry.* 65(4):409-15
Danese A, Pariante CM, Caspi A, Taylor A, Poulton R. 2007. Childhood maltreatment predicts adult inflammation in a life-course study. *Proc. Natl. Acad. Sci. U. S. A.* 104(4):1319-24
Dantzer R. 2001. Cytokine-induced sickness behavior: mechanisms and implications. *Ann. N. Y. Acad. Sci.* 933:222-34.:222-34
Dantzer R. 2005. Somatization: a psychoneuroimmune perspective. *Psychoneuroendocrinology.* 30(10):947-52
Dantzer R, Capuron L, Irwin MR, Miller AH, Ollat H et al. 2008a. Identification and treatment of symptoms associated with inflammation in medically ill patients. *Psychoneuroendocrinology.* 33(1):18-29
Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. 2008b. From inflammation to sickness and depression: when the immune system subjugates the brain. *Nat. Rev. Neurosci.* 9(1):46-56
Day SJ, Altman DG. 2000. Statistics notes: blinding in clinical trials and other studies. *BMJ.* 321(7259):504
De Becker P, Dendale P, de Meirleir K., Campine I, Vandenborne K, Hagers Y. 1998. Autonomic testing in patients with chronic fatigue syndrome. *Am. J. Med.* 105(3A):22S-6S
De Bruyne MC, Kors JA, Hoes AW, Klootwijk P, Dekker JM et al. 1999. Both decreased and increased heart rate variability on the standard 10-second electrocardiogram predict cardiac mortality in the elderly: the Rotterdam Study. *Am. J. Epidemiol.* 150(12):1282-8
De Groot M, Anderson R, Freedland KE, Clouse RE, Lustman PJ. 2001. Association of depression and diabetes complications: a meta-analysis. *Psychosom. Med.* 63(4):619-30
De Meersman RE, Stein PK. 2007. Vagal modulation and aging. *Biol. Psychol.* 74(2):165-73
Deary IJ. 1999. A taxonomy of medically unexplained symptoms. *J. Psychosom. Res.* 47(1):51-9
Deary V, Chalder T, Sharpe M. 2007. The cognitive behavioural model of medically unexplained symptoms: a theoretical and empirical review. *Clin. Psychol. Rev.* 27(7):781-97
DeCoster J. Meta-analysis Notes. 2004. Retrieved at 04-2009 from http://www.stat-help.com/notes.html.
Deeks J. 1997. Are you sure that's a standard deviation? (part 2). *Cochrane News*(11):11-2
Deeks JJ. 2002. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Stat. Med.* 21(11):1575-600
Demitrack MA, Dale JK, Straus SE, Laue L, Listwak SJ et al. 1991. Evidence for impaired activation of the hypothalamic-pituitary-adrenal axis in patients with chronic fatigue syndrome. *J. Clin. Endocrinol. Metab.* 73(6):1224-34
Denver JW, Reed SF, Porges SW. 2007. Methodological issues in the quantification of respiratory sinus arrhythmia. *Biol. Psychol.* 74(2):286-94
Derijk RH, de Kloet ER. 2008. Corticosteroid receptor polymorphisms: determinants of vulnerability and resilience. *Eur. J. Pharmacol.* 583(2-3):303-11
Derogatis LR, Lipman RS, Covi L. 1973. SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacol. Bull.* 9(1):13-28
DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. *Control. Clin. Trials.* 7(3):177-88

Dessein PH, Shipton EA, Joffe BI, Hadebe DP, Stanwix AE, Van der Merwe BA. 1999. Hyposecretion of adrenal androgens and the relation of serum adrenal steroids, serotonin and insulin-like growth factor-1 to clinical features in women with fibromyalgia. *Pain.* 83(2):313-9
Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbe KA. 1992. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J. Clin. Epidemiol.* 45(3):255-65
Di Giorgio A, Hudson M, Jerjes W, Cleare AJ. 2005. 24-hour pituitary and adrenal hormone profiles in chronic fatigue syndrome. *Psychosom. Med.* 67(3):433-40
Dickens C, McGowan L, Clark-Carter D, Creed F. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. 2002. *Psychosom. Med.* 64(1):52-60
Dickens C, McGowan L, Dale S. Impact of depression on experimental pain perception: a systematic review of the literature with meta-analysis.2003. *Psychosom. Med.* 65(3):369-75.
Dickerson SS, Kemeny ME. 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol. Bull.* 130(3):355-91
Dickhaus B, Mayer EA, Firooz N, Stains J, Conde F et al. 2003. Irritable bowel syndrome patients show enhanced modulation of visceral perception by auditory stress. *Am. J. Gastroenterol.* 98(1):135-43
Dickinson WP, Dickinson LM, deGruy FV, Candib LM, Main DS et al. 2003. The somatization in primary care study: a tale of three diagnoses. *Gen. Hosp. Psychiatry.* 25(1):1-7
Dimsdale JE, Dantzer R. 2007. A biological substrate for somatoform disorders: importance of pathophysiology. *Psychosom. Med.* 69(9):850-4
Dinan TG, Majeed T, Lavelle E, Scott LV, Berti C, Behan P. 1997. Blunted serotonin-mediated activation of the hypothalamic-pituitary-adrenal axis in chronic fatigue syndrome. *Psychoneuroendocrinology.* 22(4):261-7
Dinan TG, Quigley EM, Ahmed SM, Scully P, O'Brien S et al. 2006. Hypothalamic-pituitary-gut axis dysregulation in irritable bowel syndrome: plasma cytokines as a potential biomarker? *Gastroenterology.* 130(2):304-11
Dishman RK, Nakamura Y, Garcia ME, Thompson RW, Dunn AL, Blair SN. 2000. Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *Int. J. Psychophysiol.* 37(2):121-33
Dobrek L, Friediger J, Furgala A, Thor PJ. 2006. Autonomic nervous system activity in IBS patients estimated by heart rate variability (HRV). *Przegl. Lek.* 63(9):743-7
DuBois DF, DuBois EF. 1916. A formula to estimate the approximate surface area if height and weight are known. *Ann. Intern. Med.* 17:863-71
Duprez DA, De Buyzere ML, Drieghe B, Vanhaverbeke F, Taes Y et al. 1998. Long- and short-term blood pressure and RR-interval variability and psychosomatic distress in chronic fatigue syndrome. *Clin. Sci. (Lond).* 94(1):57-63
Duval S, Tweedie R. 2000a. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *J. Am. Stat. Assoc.* 95(449):89-98
Duval S, Tweedie R. 2000b. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics.* 56(2):455-63
Edwards P, Clarke M, DiGuiseppi C, Pratap S, Roberts I, Wentz R. 2002. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat. Med.* 21(11):1635-40
Edwards S, Clow A, Evans P, Hucklebridge F. 2001. Exploration of the awakening cortisol response in relation to diurnal cortisol secretory activity. *Life Sci.* 68(18):2093-103
Egger M, Davey SG, Schneider M, Minder C. 1997a. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* 315(7109):629-34
Egger M, Schneider M, Davey SG. 1998a. Spurious precision? Meta-analysis of observational studies. *BMJ.* 316(7125):140-4
Egger M, Smith GD. 1997d. Meta-Analysis. Potentials and promise. *BMJ.* 315(7119):1371-4
Egger M, Smith GD, Phillips AN. 1997b. Meta-analysis: principles and procedures. *BMJ.* 315(7121):1533-7
Egger M, Stern JA, Smith GD. 1998b. Meta-Analysis Software. *BMJ.* 316(7126):221-5
Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G. 1997c. Language bias in randomised controlled trials published in English and German. *Lancet.* 350(9074):326-9
Ehlert U, Gaab J, Heinrichs M. 2001. Psychoneuroendocrinological contributions to the etiology of depression, posttraumatic stress disorder, and stress-related bodily disorders: the role of the hypothalamus-pituitary-adrenal axis. *Biol. Psychol.* 57(1-3):141-52

Ehlert U, Nater UM, Bohmelt A. 2005. High and low unstimulated salivary cortisol levels correspond to different symptoms of functional gastrointestinal disorders. *J. Psychosom. Res.* 59(1):7-10
Elsenbruch S, Harnish MJ, Orr WC. 1999. Subjective and objective sleep quality in irritable bowel syndrome. *Am. J. Gastroenterol.* 94(9):2447-52
Elsenbruch S, Holtmann G, Oezcan D, Lysson A, Janssen O et al. 2004. Are there alterations of neuroendocrine and cellular immune responses to nutrients in women with irritable bowel syndrome? *Am. J. Gastroenterol.* 99(4):703-10
Elsenbruch S, Lovallo WR, Orr WC. 2001a. Psychological and physiological responses to postprandial mental stress in women with the irritable bowel syndrome. *Psychosom. Med.* 63(5):805-13
Elsenbruch S, Lucas A, Holtmann G, Haag S, Gerken G et al. 2006. Public speaking stress-induced neuroendocrine responses and circulating immune cell redistribution in irritable bowel syndrome. *Am. J. Gastroenterol.* 101(10):2300-7
Elsenbruch S, Orr WC. 2001b. Diarrhea- and constipation-predominant IBS patients differ in postprandial autonomic and cortisol responses. *Am. J. Gastroenterol.* 96(2):460-6
Elsenbruch S, Thompson JJ, Hamish MJ, Exton MS, Orr WC. 2002. Behavioral and physiological sleep characteristics in women with irritable bowel syndrome. *Am. J. Gastroenterol.* 97(9):2306-14
Elzinga BM, Roelofs K, Tollenaar MS, Bakvis P, van PJ, Spinhoven P. 2008. Diminished cortisol responses to psychosocial stress associated with lifetime adverse events a study among healthy young subjects. *Psychoneuroendocrinology.* 33(2):227-37
Ergun U, Demirci M, Nurlu G, Komurcu F. 2008. Power spectral analysis of heart rate variability: normal values of subjects over 60 years old. *Int. J. Neurosci.* 118(8):1165-73
Eriksen HR, Ursin H. 2004. Subjective health complaints, sensitization, and sustained cognitive activation (stress). *J. Psychosom. Res.* 56(4):445-8
Eriksson EM, Andren KI, Eriksson HT, Kurlberg GK. 2008. Irritable bowel syndrome subtypes differ in body awareness, psychological symptoms and biochemical stress markers. *World J. Gastroenterol.* 14(31):4889-96
Eriksson EM, Moller IE, Soderberg RH, Eriksson HT, Kurlberg GK. 2007. Body awareness therapy: a new strategy for relief of symptoms in irritable bowel syndrome patients. *World J. Gastroenterol.* 13(23):3206-14
Escobar JI, Burnam MA, Karno M, Forsythe A, Golding JM. 1987. Somatization in the community. *Arch. Gen. Psychiatry* 44(8):713-8
Escobar JI, Rubio-Stipec M, Canino G, Karno M. 1989. Somatic symptom index (SSI): a new and abridged somatization construct. Prevalence and epidemiological correlates in two large community samples. *J. Nerv. Ment. Dis.* 177(3):140-6
Escobar JI, Waitzkin H, Silver RC, Gara M, Holman A. 1998. Abridged somatization: a study in primary care. *Psychosom. Med.* 60(4):466-72
Evans S. 1996. Statistician's comment on: Misleading meta-analysis. "Fail safe N" is a useful mathematical measure of the stability of results. *BMJ.* 312(7023):125
Fabian LA, McGuire L, Page GG, Goodin BR, Edwards RR, Haythornthwaite J. 2009. The association of the cortisol awakening response with experimental pain ratings. *Psychoneuroendocrinology.* 34(8):1247-51
Fava GA, Sonino N. 2009. Psychosomatic Assessment. *Psychother. Psychosom.* 78(6):333-41
Fietta P, Fietta P. 2006. Counterbalance between leptin and cortisol may be associated with fibromyalgia. *Psychiatry Clin. Neurosci.* 60(4):529
Fink P. 1996. Somatization--beyond symptom count. *J. Psychosom. Res.* 40(1):7-10
Fink P, Sorensen L, Engberg M, Holm M, Munk-Jorgensen P. 1999. Somatization in primary care. Prevalence, health care utilization, and general practitioner recognition. *Psychosomatics.* 40(4):330-8
Fink P, Toft T, Hansen MS, Ornbol E, Olesen F. 2007. Symptoms and syndromes of bodily distress: an exploratory study of 978 internal medical, neurological, and primary care patients. *Psychosom. Med.* 69(1):30-9
Finset A, Graugaard PK, Holgersen K. 2006. Salivary cortisol response after a medical interview: the impact of physician communication behaviour, depressed affect and alexithymia. *Patient. Educ. Couns.* 60(2):115-24
Fleiss JL. 1993. The statistical basis of meta-analysis. *Stat. Methods Med. Res.* 2(2):121-45
Folks DG. 2004. The interface of psychiatry and irritable bowel syndrome. *Curr. Psychiatry Rep.* 6(3):210-5

Freedland KE, Reese RL, Steinmeyer BC. 2009. Multivariable models in biobehavioral research. *Psychosom. Med.* 71(2):205-16
Friederich HC, Schellberg D, Mueller K, Bieber C, Zipfel S, Eich W. 2005. Stress and autonomic dysregulation in patients with fibromyalgia syndrome. *Schmerz.* 19(3):185-94
Friedman BH. 2007. An autonomic flexibility-neurovisceral integration model of anxiety and cardiac vagal tone. *Biol. Psychol.* 74(2):185-99
Fries E. 2008. Hypocortisolemic Disorders. In *Stress. The Brain-Body Connection. Key Issues in Mental Health*, ed. DH Helhammer, J Helhammer, 5:60-77. Basel, Karger.
Fries E, Hesse J, Hellhammer J, Hellhammer DH. 2005. A new view on hypocortisolism. *Psychoneuroendocrinology.* 30(10):1010-6
Frostholm L, Oernboel E, Christensen KS, Toft T, Olesen F et al. 2007. Do illness perceptions predict health outcomes in primary care patients? A 2-year follow-up study. *J. Psychosom. Res.* 62(2):129-38
Fuite J, Vernon SD, Broderick G. 2008. Neuroendocrine and immune network re-modeling in chronic fatigue syndrome: an exploratory analysis. *Genomics.* 92(6):393-9
Fukudo S. 2007. Role of corticotropin-releasing hormone in irritable bowel syndrome and intestinal inflammation. *J. Gastroenterol.* 42 Suppl 17:48-51.:48-51
Fukudo S, Kanazawa M, Kano M, Sagami Y, Endo Y et al. 2002. Exaggerated motility of the descending colon with repetitive distention of the sigmoid colon in patients with irritable bowel syndrome. *J. Gastroenterol.* 37 Suppl 14:145-50
Fukudo S, Nomura T, Hongo M. 1998. Impact of corticotropin-releasing hormone on gastrointestinal motility and adrenocorticotropic hormone in normal controls and patients with irritable bowel syndrome. *Gut.* 42(6):845-9
Furlan R, Colombo S, Perego F, Atzeni F, Diana A et al. 2005. Abnormalities of cardiovascular neural control and reduced orthostatic tolerance in patients with primary fibromyalgia. *J. Rheumatol.* 32(9):1787-93
Gaab J, Engert V, Heitz V, Schad T, Schurmeyer TH, Ehlert U. 2004. Associations between neuroendocrine responses to the Insulin Tolerance Test and patient characteristics in chronic fatigue syndrome. *J. Psychosom. Res.* 56(4):419-24
Gaab J, Huster D, Peisen R, Engert V, Heitz V et al. 2003. Assessment of cortisol response with low-dose and high-dose ACTH in patients with chronic fatigue syndrome and healthy comparison subjects. *Psychosomatics.* 44(2):113-9
Gaab J, Huster D, Peisen R, Engert V, Heitz V et al. 2002a. Hypothalamic-pituitary-adrenal axis reactivity in chronic fatigue syndrome and health under psychological, physiological, and pharmacological stimulation. *Psychosom. Med.* 64(6):951-62
Gaab J, Huster D, Peisen R, Engert V, Schad T et al. 2002b. Low-dose dexamethasone suppression test in chronic fatigue syndrome and health. *Psychosom. Med.* 64(2):311-8
Gaab J, Rohleder N, Heitz V, Engert V, Schad T et al. 2005. Stress-induced changes in LPS-induced pro-inflammatory cytokine production in chronic fatigue syndrome. *Psychoneuroendocrinology.* 30(2):188-98
Gara MA, Silver RC, Escobar JI, Holman A, Waitzkin H. 1998. A hierarchical classes analysis (HICLAS) of primary care patients with medically unexplained somatic symptoms. *Psychiatry Res.* 19;81(1):77-86
Garg AX, Hackam D, Tonelli M. 2008. Systematic review and meta-analysis: when one study is just not enough. *Clin. J. Am. Soc. Nephrol.* 3(1):253-60
Geenen R, Jacobs JW, Bijlsma JW. 2002. Evaluation and management of endocrine dysfunction in fibromyalgia. *Rheum. Dis. Clin. North Am.* 28(2):389-404
Genuth SM. 1998. The hypothalamus and pituitary gland. In *Physiology, Fourth Edition*, ed. RM Berne, RL Levy, 49:872-909. Missouri, Mosby.
Ghione S. 1996. Hypertension-associated hypalgesia. Evidence in experimental animals and humans, pathophysiological mechanisms, and potential clinical consequences. *Hypertension.* 28(3):494-504
Gibson CA, Bailey BW, Carper MJ, Lecheminant JD, Kirk EP et al. 2006. Author contacts for retrieval of data for a meta-analysis on exercise and diet restriction. *Int. J. Technol. Assess. Health Care.* 22(2):267-70
Gillespie NA, Zhu G, Heath AC, Hickie IB, Martin NG. 2000. The genetic aetiology of somatic distress. *Psychol. Med.* 30(5):1051-61

Giske L, Vollestad NK, Mengshoel AM, Jensen J, Knardahl S, Roe C. 2008. Attenuated adrenergic responses to exercise in women with fibromyalgia--a controlled study. *Eur. J. Pain.* 12(3):351-60

Glass JM, Lyden AK, Petzke F, Stein P, Whalen G et al. 2004. The effect of brief exercise cessation on pain, fatigue, and mood symptom development in healthy, fit individuals. *J. Psychosom. Res.* 57(4):391-8

Glei DA, Goldman N, Chuang YL, Weinstein M. 2007. Do chronic stressors lead to physiological dysregulation? Testing the theory of allostatic load. *Psychosom. Med.* 69(8):769-76

Goertzel BN, Pennachin C, de Souza CL, Maloney EM, Jones JF, Gurbaxani B. 2006. Allostatic load is associated with symptoms in chronic fatigue syndrome patients. *Pharmacogenomics.* 7(3):485-94

Gold PW, Licinio J, Wong ML, Chrousos GP. 1995. Corticotropin releasing hormone in the pathophysiology of melancholic and atypical depression and in the mechanism of action of antidepressant drugs. *Ann. N. Y. Acad. Sci.* 771:716-29

Goodman LA. 1974a. Analysis of Systems of Qualitative Variables When Some of Variables Are Unobservable. 1. Modified Latent Structure Approach. *Am. J. Sociology* 79(5):1179-259

Goodman LA. 1974b. Exploratory Latent Structure-Analysis Using Both Identifiable and Unidentifiable Models. *Biometrika.* 61(2):215-31

Gotzsche PC, Hrobjartsson A, Maric K, Tendal B. 2007. Data extraction errors in meta-analyses that use standardized mean differences. *JAMA.* 298(4):430-7

Graham JE, Robles TF, Kiecolt-Glaser JK, Malarkey WB, Bissell MG, Glaser R. 2006. Hostility and pain are related to inflammation in older adults. *Brain Behav. Immun.* 20(4):389-400

Granger DA, Hibel LC, Fortunato CK, Kapelewski CH. 2009. Medication effects on salivary cortisol: Tactics and strategy to minimize impact in behavioral and developmental science. *Psychoneuroendocrinology.* 2009;34(10):1437-48

Greenland S, O'Rourke K. 2008. Meta-Analysis. In *Modern Epidemiology*, ed. KJ Rothman, S Greenland, TJ Lash, 33:652-682. Philadelphia, Lippincott Williams & Wilkins.

Griep EN. 2000. Difference in hypothalamic-pituitary-adrenal response pattern in fibromyalgia and chronic fatigue syndrome. In: *Dissertation, Fibromyalgia syndrome. Neuroendocrine Perspectives with Clinical Implications*, 7:131-174.

Griep EN, Boersma JW, de Kloet ER. 1993a. Altered reactivity of the hypothalamic-pituitary-adrenal axis in the primary fibromyalgia syndrome. *J. Rheumatol.* 20(3):469-74

Griep EN, Boersma JW, de Kloet ER. 1993b. Evidence for neuroendocrine disturbance following physical exercise in primary fibromyalgia syndrome. *J. Musculoskelet. Pain* 1(3-4):217-22

Griep EN, Boersma JW, de Kloet ER. 1994. Pituitary release of growth hormone and prolactin in the primary fibromyalgia syndrome. *J. Rheumatol.* 21(11):2125-30

Griep EN, Boersma JW, Lentjes EG, Prins AP, van der Korst JK, de Kloet ER. 1998. Function of the hypothalamic-pituitary-adrenal axis in patients with fibromyalgia and low back pain. *J. Rheumatol.* 25(7):1374-81

Gupta A, Silman AJ. 2004. Psychological stress and fibromyalgia: a review of the evidence suggesting a neuroendocrine link. *Arthritis Res. Ther.* 6(3):98-106

Gur A, Cevik R, Nas K, Colpan L, Sarac S. 2004a. Cortisol and hypothalamic-pituitary-gonadal axis hormones in follicular-phase women with fibromyalgia and chronic fatigue syndrome and effect of depressive symptoms on these hormones. *Arthritis Res. Ther.* 6(3):R232-R238

Gur A, Cevik R, Sarac AJ, Colpan L, Em S. 2004b. Hypothalamic-pituitary-gonadal axis and cortisol in young women with primary fibromyalgia: the potential roles of depression, fatigue, and sleep disturbance in the occurrence of hypocortisolism. *Ann. Rheum. Dis.* 63(11):1504-6

Gur A, Oktayoglu P. 2008. Status of immune mediators in fibromyalgia. *Curr. Pain Headache Rep.* 12(3):175-81

Gursel Y, Ergin S, Ulus Y, Erdogan MF, Yalcin P, Evcik D. 2001. Hormonal responses to exercise stress test in patients with fibromyalgia syndrome. *Clin. Rheumatol.* 20(6):401-5

Haapaniemi TH, Pursiainen V, Korpelainen JT, Huikuri HV, Sotaniemi KA, Myllyla VV. 2001. Ambulatory ECG and analysis of heart rate variability in Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry.* 70(3):305-10

Hamilos DL, Nutter D, Gershtenson J, Redmond DP, Clementi JD et al. 1998. Core body temperature is normal in chronic fatigue syndrome. *Biol. Psychiatry.* 43(4):293-302

Hardy RJ, Thompson SG. 1998. Detecting and describing heterogeneity in meta-analysis. *Stat. Med.* 17(8):841-56

Hart BL. 1988. Biological basis of the behavior of sick animals. *Neurosci. Biobehav. Rev.* 12(2):123-37
Hartz AJ, Noyes R, Bentler SE, Damiano PC, Willard JC, Momany ET. 2000. Unexplained symptoms in primary care: perspectives of doctors and patients. *Gen. Hosp. Psychiatry.* 22(3):144-52
Harvey SB, Wadsworth M, Wessely S, Hotopf M. 2008. The relationship between prior psychiatric disorder and chronic fatigue: evidence from a national birth cohort study. *Psychol. Med.* 38(7):933-40
Hasin D, Katz H. 2007. Somatoform and substance use disorders. *Psychosom. Med.* 69(9):870-5
Haug TT, Mykletun A, Dahl AA. 2004. The association between anxiety, depression, and somatic symptoms in a large population: the HUNT-II study. *Psychosom. Med.* 66(6):845-51
Hayano J, Yamada M, Sakakibara Y, Fujinami T, Yokoyama K et al. 1990. Short- and long-term effects of cigarette smoking on heart rate variability. *Am. J. Cardiol.* 65(1):84-8
Hedges LV. 1981. Distribution theory for Glass's estimator of effect size and related estimators. *J. Educ. Stat* 6:107-28
Hedges LV, Olkin I. 1985. *Statistical methods for meta-analysis,* San Diego, CA, Academic Press.
Heim C, Ehlert U, Hellhammer DH. 2000a. The potential role of hypocortisolism in the pathophysiology of stress-related bodily disorders. *Psychoneuroendocrinology.* 25(1):1-35
Heim C, Nater UM, Maloney E, Boneva R, Jones JF, Reeves WC. 2009. Childhood trauma and risk for chronic fatigue syndrome: association with neuroendocrine dysfunction. *Arch. Gen. Psychiatry.* 66(1):72-80
Heim C, Newport DJ, Heit S, Graham YP, Wilcox M et al. 2000b. Pituitary-adrenal and autonomic responses to stress in women after sexual and physical abuse in childhood. *JAMA.* 284(5):592-7
Heim C, Wagner D, Maloney E, Papanicolaou DA, Solomon L et al. 2006. Early adverse experience and risk for chronic fatigue syndrome: results from a population-based study. *Arch. Gen. Psychiatry.* 63(11):1258-66
Heitkemper M, Jarrett M, Cain K, Shaver J, Bond E et al. 1996. Increased urine catecholamines and cortisol in women with irritable bowel syndrome. *Am. J. Gastroenterol.* 91(5):906-13
Heitkemper M, Jarrett M, Cain K, Shaver J, Bond E et al. 2001. Effect of sexual and physical abuse on symptom experiences in women with irritable bowel syndrome. *Nursing Res.* 50(1):15-23
Hendriks AA, Ormel J, van de Willige G. 1990. Langdurige moeilijkheden gemeten volgens zelfbeoordelingsvragenlijst en semi-gestructureerd interview: een theoretische en empirische vergelijking. *Gedrag & Gezondheid.* 18(6):273-83
Henningsen P, Zimmermann T, Sattel H. 2003. Medically unexplained physical symptoms, anxiety, and depression: a meta-analytic review. *Psychosom. Med.* 65(4):528-33
Henningsen P, Zipfel S, Herzog W. 2007. Management of functional somatic syndromes. *Lancet.* 369(9565):946-55
Herbison P, Hay-Smith J, Gillespie WJ. 2006. Adjustment of meta-analyses on the basis of quality scores should be abandoned. *J. Clin. Epidemiol.* 59(12):1249-56
Higgins JP, Thompson S, Deeks J, Altman D. 2002a. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *J. Health Serv. Res. Policy.* 7(1):51-61
Higgins JP, Green S. 2008. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.1.*
Higgins JP, Thompson SG. 2002b. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* 21(11):1539-58
Higgins JP, Thompson SG. 2004. Controlling the risk of spurious findings from meta-regression. *Stat. Med.* 23(11):1663-82
Higgins JP, White IR, Anzures-Cabrera J. 2008. Meta-analysis of skewed data: Combining results reported on log-transformed or raw scales. *Stat. Med.* 20;27(29):6072-92
Hiller W, Cuntz U, Rief W, Fichter MM. 2001. Searching for a gastrointestinal subgroup within the somatoform disorders. *Psychosomatics.* 42(1):14-20
Hiller W, Rief W, Brahler E. 2006. Somatization in the population: from mild bodily misperceptions to disabling symptoms. *Soc. Psychiatry Psychiatr. Epidemiol.* 41(9):704-12
Holsboer F, Barden N. 1996. Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocr. Rev.* 17(2):187-205
Hotopf M, Mayou R, Wadsworth M, Wessely S. 1999. Childhood risk factors for adults with medically unexplained symptoms: results from a national birth cohort study. *Am. J. Psychiatry.* 156(11):1796-800

Houtveen JH, Kavelaars A, Heijnen CJ, van Doornen LJ. 2007a. Heterogeneous medically unexplained symptoms and immune function. *Brain Behav. Immun.* 21(8):1075-82
Houtveen JH, Rietveld S, de Geus EJ. 2003. Exaggerated perception of normal physiological responses to stress and hypercapnia in young women with numerous functional somatic symptoms. *J. Psychosom. Res.* 55(6):481-90
Houtveen JH, van Doornen LJ. 2007b. Medically unexplained symptoms and between-group differences in 24-h ambulatory recording of stress physiology. *Biol. Psychol.* 76(3):239-49
Houtveen JH, van Doornen LJ. 2008. Negative affect and 24-hour ambulatory physiological recordings as predictors of spontaneous improvement of medically unexplained symptoms. *Scand. J. Psychol.* 49(6):591-601
Howren MB, Lamkin DM, Suls J. 2009. Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis. *Psychosom. Med.* 71(2):171-86
Hu L, Bentler PM. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Eq. Modeling* 6(1):1-55
Hudson M, Cleare AJ. 1999. The 1microg short Synacthen test in chronic fatigue syndrome. *Clin. Endocrinol. (Oxf).* 51(5):625-30
Hutchinson WL, Koenig W, Frohlich M, Sund M, Lowe GD, Pepys MB. 2000. Immunoradiometric assay of circulating C-reactive protein: age-related values in the adult general population. *Clin. Chem.* 46(7):934-8
Ihlebaek C, Eriksen HR, Ursin H. 2002. Prevalence of subjective health complaints (SHC) in Norway. *Scand. J. Public Health.* 30(1):20-9
Inder WJ, Prickett TC, Mulder RT. 2005. Normal opioid tone and hypothalamic-pituitary-adrenal axis function in chronic fatigue syndrome despite marked functional impairment. *Clin. Endocrinol. (Oxf).* 62(3):343-8
Ioannidis JP. 2005. Why most published research findings are false. *PLoS. Med.* 2(8):e124
Ioannidis JP. 2008. Interpretation of tests of heterogeneity and bias in meta-analysis. *J. Eval. Clin. Pract.* 14(5):951-7
Iversen A, Chalder T, Wessely S. 2007. Gulf War Illness: lessons from medically unexplained symptoms. *Clin. Psychol. Rev.* 27(7):842-54
Izgi HB, Gokce C, Calis M, Turan T, Kirnap M et al. 2005. Investigation of the hypothalamopituitary-adrenal axis by low-dose (1 lg) adrenocorticotrophic hormone test and metyrapone test in patients with chronic fatigue syndrome. *Endocrinologist.* 15(2):89-92
Izquierdo-Alvarez S, Bocos-Terraz JP, Bancalero-Flores JL, Pavon-Romero L, Serrano-Ostariz E, de Miquel CA. 2008. Is there an association between fibromyalgia and below-normal levels of urinary cortisol? *BMC. Res. Notes.* 1:134.:134
Jackson J, Fiddler M, Kapur N, Wells A, Tomenson B, Creed F. 2006. Number of bodily symptoms predicts outcome more accurately than health anxiety in patients attending neurology, cardiology, and gastroenterology clinics. *J. Psychosom. Res.* 60(4):357-63
Jackson JL, Kroenke K. 2008. Prevalence, impact, and prognosis of multisomatoform disorder in primary care: a 5-year follow-up study. *Psychosom. Med.* 70(4):430-4
Jerjes WK, Cleare AJ, Wessely S, Wood PJ, Taylor NF. 2005. Diurnal patterns of salivary cortisol and cortisone output in chronic fatigue syndrome. *J. Affect. Disord.* 87(2-3):299-304
Jerjes WK, Peters TJ, Taylor NF, Wood PJ, Wessely S, Cleare AJ. 2006a. Diurnal excretion of urinary cortisol, cortisone, and cortisol metabolites in chronic fatigue syndrome. *J. Psychosom. Res.* 60(2):145-53
Jerjes WK, Taylor NF, Peters TJ, Wessely S, Cleare AJ. 2006b. Urinary cortisol and cortisol metabolite excretion in chronic fatigue syndrome. *Psychosom. Med.* 68(4):578-82
Jerjes WK, Taylor NF, Wood PJ, Cleare AJ. 2007. Enhanced feedback sensitivity to prednisolone in chronic fatigue syndrome. *Psychoneuroendocrinology.* 32(2):192-8
Jindal RD, Vasko RC, Jr., Jennings JR, Fasiczka AL, Thase ME, Reynolds CF, III. 2008. Heart rate variability in depressed elderly. *Am. J. Geriatr. Psychiatry.* 16(11):861-6
John U, Meyer C, Rumpf HJ, Hapke U. 2004. Smoking, nicotine dependence and psychiatric comorbidity--a population-based study including smoking cessation after three years. *Drug Alcohol Depend.* 76(3):287-95
Jones DR. 1992. Meta-analysis of observational epidemiological studies: a review. *J. R. Soc. Med.* 85(3):165-8
Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. 2002. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int. J. Epidemiol.* 31(1):115-23

Juni P, Witschi A, Bloch R, Egger M. 1999. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA.* 282(11):1054-60
Kanner AD, Coyne JC, Schaefer C, Lazarus RS. 1981. Comparison of two modes of stress measurement: daily hassles and uplifts versus major life events. *J. Behav. Med.* 4(1):1-39
Karakaya O, Barutcu I, Kaya D, Esen AM, Saglam M et al. 2007. Acute effect of cigarette smoking on heart rate variability. *Angiology.* 58(5):620-4
Karling P, Nyhlin H, Wiklund U, Sjoberg M, Olofsson BO, Bjerle P. 1998. Spectral analysis of heart rate variability in patients with irritable bowel syndrome. *Scand. J. Gastroenterol.* 33(6):572-6
Kato K, Sullivan PF, Evengard B, Pedersen NL. 2006. Premorbid predictors of chronic fatigue. *Arch. Gen. Psychiatry.* 63(11):1267-72
Kato K, Sullivan PF, Evengard B, Pedersen NL. 2009. A population-based twin study of functional somatic syndromes. *Psychol. Med.* 39(3):497-505
Katon W, Lin E, Von KM, Russo J, Lipscomb P, Bush T. 1991. Somatization: a spectrum of severity. *Am. J. Psychiatry.* 148(1):34-40
Katon WJ, Walker EA. 1998. Medically unexplained symptoms in primary care. *J. Clin. Psychiatry.* 59 Suppl 20:15-21
Kaufmann I, Schelling G, Eisner C, Richter HP, Krauseneck T et al. 2008. Anandamide and neutrophil function in patients with fibromyalgia. *Psychoneuroendocrinology.* 33(5):676-85
Kiecolt-Glaser JK, McGuire L, Robles TF, Glaser R. 2002. Psychoneuroimmunology and psychosomatic medicine: back to the future. *Psychosom. Med.* 64(1):15-28
Kilkens TO, Honig A, Fekkes D, Brummer RJ. 2005. The effects of an acute serotonergic challenge on brain-gut responses in irritable bowel syndrome patients and controls. *Aliment. Pharmacol. Ther.* 22(9):865-74
Kindt S, Van OL, Broekaert D, Kasran A, Ceuppens JL et al. 2009. Immune dysfunction in patients with functional gastrointestinal disorders. *Neurogastroenterol. Motil.* 21(4):389-98
Kingma EM, Tak LM, Huisman M, Rosmalen JG. 2009. Intelligence is negatively associated with the number of functional somatic symptoms. *J Epidemiol. Community Health.* 63(11):900-5.
Kirmayer LJ, Looper KJ. 2006. Abnormal illness behaviour: physiological, psychological and social dimensions of coping with distress. *Curr. Opin. Psychiatry.* 19(1):54-60
Kirnap M, Colak R, Eser C, Ozsoy O, Tutus A, Kelestimur F. 2001. A comparison between low-dose (1 microg), standard-dose (250 microg) ACTH stimulation tests and insulin tolerance test in the evaluation of hypothalamo-pituitary-adrenal axis in primary fibromyalgia syndrome. *Clin. Endocrinol. (Oxf).* 55(4):455-9
Kirschbaum C, Hellhammer DH. 1994. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology.* 19(4):313-33
Kirschbaum C, Pirke KM, Hellhammer DH. 1993. The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology.* 28(1-2):76-81
Kisely S, Goldberg D, Simon G. 1997. A comparison between somatic symptoms with and without clear organic cause: results of an international study. *Psychol. Med.* 27(5):1011-9
Klerman EB, Goldenberg DL, Brown EN, Maliszewski AM, Adler GK. 2001. Circadian rhythms of women with fibromyalgia. *J. Clin. Endocrinol. Metab* 86(3):1034-9
Klimas NG, Koneru AO. 2007. Chronic fatigue syndrome: inflammation, immune function, and neuroendocrine interactions. *Curr. Rheumatol. Rep.* 9(6):482-7
Klingmann PO, Kugler I, Steffke TS, Bellingrath S, Kudielka BM, Hellhammer DH. 2008. Sex-specific prenatal programming: a risk for fibromyalgia? *Ann. N. Y. Acad. Sci.* 1148:446-55
Kon H, Nagano M, Tanaka F, Satoh K, Segawa T, Nakamura M. 2006. Association of decreased variation of R-R interval and elevated serum C-reactive protein level in a general population in Japan. *Int. Heart J.* 47(6):867-76
Kop WJ, Lyden A, Berlin AA, Ambrose K, Olsen C et al. 2005. Ambulatory monitoring of physical activity and symptoms in fibromyalgia and chronic fatigue syndrome. *Arthritis Rheum.* 52(1):296-303
Korszun A, Sackett-Lundeen L, Papadopoulos E, Brucksch C, Masterson L et al. 1999. Melatonin levels in women with fibromyalgia and chronic fatigue syndrome. *J. Rheumatol.* 26(12):2675-80
Koskinen P, Virolainen J, Kupari M. 1994. Acute alcohol intake decreases short-term heart rate variability in healthy subjects. *Clin. Sci. (Lond).* 87(2):225-30
Kraemer HC, Kazdin AE, Offord DR, Kessler RC, Jensen PS, Kupfer DJ. 1997. Coming to terms with the terms of risk. *Arch. Gen. Psychiatry.* 54(4):337-43

Kraemer HC, Stice E, Kazdin A, Offord D, Kupfer D. 2001. How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *Am. J. Psychiatry.* 158(6):848-56
Kroenke K. 2001. Studying symptoms: sampling and measurement issues. *Ann. Intern. Med.* 134(9 Pt 2):844-53
Kroenke K, Mangelsdorff AD. 1989. Common symptoms in ambulatory care: incidence, evaluation, therapy, and outcome. *Am. J. Med.* 86(3):262-6
Kroenke K, Price RK. 1993. Symptoms in the community. Prevalence, classification, and psychiatric comorbidity. *Arch. Intern. Med.* 153(21):2474-80
Kroenke K, Rosmalen JG. 2006. Symptoms, syndromes, and the value of psychiatric diagnostics in patients who have functional somatic disorders. *Med. Clin. North Am.* 90(4):603-26
Kroenke K, Sharpe M, Sykes R. 2007. Revising the classification of somatoform disorders: key questions and preliminary recommendations. *Psychosomatics.* 48(4):277-85
Kroenke K, Spitzer RL. 1998a. Gender differences in the reporting of physical and somatoform symptoms. *Psychosom. Med.* 60(2):150-5
Kroenke K, Spitzer RL, deGruy FV, III, Swindle R. 1998b. A symptom checklist to screen for somatoform disorders in primary care. *Psychosomatics.* 39(3):263-72
Kroenke K, Spitzer RL, deGruy FV, III, Hahn SR, Linzer M et al. 1997. Multisomatoform disorder. An alternative to undifferentiated somatoform disorder for the somatizing patient in primary care. *Arch. Gen. Psychiatry.* 54(4):352-8
Kuch B, Hense HW, Sinnreich R, Kark JD, von EA et al. 2001. Determinants of short-period heart rate variability in the general population. *Cardiology.* 95(3):131-8
Kuh D, Ben-Shlomo Y, Lynch J, Hallqvist J, Power C. 2003. Life course epidemiology. *J Epidemiol. Community Health.* 57(10):778-83
Kumari M, Badrick E, Chandola T, Adam EK, Stafford M et al. 2009. Cortisol secretion and fatigue: Associations in a community based cohort. *Psychoneuroendocrinology.* 2009;34(10):1476-85
Kuo HK, Yen CJ, Chen JH, Yu YH, Bean JF. 2007. Association of cardiorespiratory fitness and levels of C-reactive protein: data from the National Health and Nutrition Examination Survey 1999-2002. *Int. J. Cardiol.* 114(1):28-33
Kurina LM, Weiss LA, Graves SW, Parry R, Williams GH et al. 2005. Sex differences in the genetic basis of morning serum cortisol levels: genome-wide screen identifies two novel loci specific to women. *J. Clin. Endocrinol. Metab.* 90(8):4747-52
Kushner I, Rzewnicki D, Samols D. 2006. What does minor elevation of C-reactive protein signify? *Am. J. Med.* 119(2):17-28
Kushnir MM, Rockwood AL, Nelson GJ, Terry AH, Meikle AW. 2003. Liquid chromatography-tandem mass spectrometry analysis of urinary free cortisol. *Clin. Chem.* 49(6 Pt 1):965-7
Laederach-Hofmann K, Ruddel H, Mussgay L. 2008. Pathological baroreceptor sensitivity in patients suffering from somatization disorders: Do they correlate with symptoms? *Biol. Psychol.* 79(2):243-9
Lakoski SG, Cushman M, Criqui M, Rundek T, Blumenthal RS et al. 2006. Gender and C-reactive protein: data from the Multiethnic Study of Atherosclerosis (MESA) cohort. *Am. Heart J.* 152(3):593-8
Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics.* 33(1):159-74
Lariviere WR, Melzack R. 2000. The role of corticotropin-releasing factor in pain and analgesia. *Pain.* 84(1):1-12
Lazarus RS, Folkman S. 1984. *Stress, appraisal, and coping,* New York, Springer.
Lee W, Bindman J, Ford T, Glozier N, Moran P et al. 2007. Bias in psychiatric case-control studies: literature survey. *Br. J. Psychiatry.* 190:204-9
Leiknes KA, Finset A, Moum T, Sandanger I. 2006. Methodological issues concerning lifetime medically unexplained and medically explained symptoms of the Composite International Diagnostic Interview: a prospective 11-year follow-up study. *J. Psychosom. Res.* 61(2):169-79
Lentjes EG, Griep EN, Boersma JW, Romijn FP, de Kloet ER. 1997. Glucocorticoid receptors, fibromyalgia and low back pain. *Psychoneuroendocrinology.* 22(8):603-14
Levine A, Zagoory-Sharon O, Feldman R, Lewis JG, Weller A. 2007. Measuring cortisol in human psychobiological studies. *Physiol Behav.* 90(1):43-53

Licht CM, de Geus EJ, van DR, Penninx BW. 2009. Association between anxiety disorders and heart rate variability in The Netherlands Study of Depression and Anxiety (NESDA). *Psychosom. Med.* 71(5):508-18
Licht CM, de Geus EJ, Zitman FG, Hoogendijk WJ, van DR, Penninx BW. 2008. Association Between Major Depressive Disorder and Heart Rate Variability in the Netherlands Study of Depression and Anxiety (NESDA). *Arch. Gen. Psychiatry.* 65(12):1358-67
Licinio J, Wong ML, Gold PW. 1996. The hypothalamic-pituitary-adrenal axis in anorexia nervosa. *Psychiatry Res.* 62(1):75-83
Liebregts T, Adam B, Bredack C, Roth A, Heinzel S et al. 2007. Immune activation in patients with irritable bowel syndrome. *Gastroenterology.* 132(3):913-20
Light KC, Bragdon EE, Grewen KM, Brownley KA, Girdler SS, Maixner W. 2009. Adrenergic dysregulation and pain with and without acute beta-blockade in women with fibromyalgia and temporomandibular disorder. *J. Pain.* 10(5):542-552.
Lindqvist D, Isaksson A, Traskman-Bendz L, Brundin L. 2008. Salivary cortisol and suicidal behavior--a follow-up study. *Psychoneuroendocrinology.* 33(8):1061-8
Lipowski ZJ. 1988. Somatization: the concept and its clinical application. *Am. J. Psychiatry.* 145(11):1358-68
Lipsey MW, Wilson DB. 2001. *Practical Meta-Analysis,* Sage Publication Ltd.
Liu G, Clark MR, Eaton WW. 1997. Structural factor analyses for medically unexplained somatic symptoms of somatization disorder in the Epidemiologic Catchment Area study. *Psychol. Med.* 27(3):617-26
Liukkonen T, Silvennoinen-Kassinen S, Jokelainen J, Rasanen P, Leinonen M et al. 2006. The association between C-reactive protein levels and depression: Results from the northern Finland 1966 birth cohort study. *Biol. Psychiatry.* 60(8):825-30
Loucks EB, Juster RP, Pruessner JC. 2008. Neuroendocrine biomarkers, allostatic load, and the challenge of measurement: A commentary on Gersten. *Soc. Sci. Med.* 66(3):525-30
Lowe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. 2008. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. *Gen. Hosp. Psychiatry.* 30(3):191-9
Lu CL, Zou X, Orr WC, Chen JD. 1999. Postprandial changes of sympathovagal balance measured by heart rate variability. *Dig. Dis. Sci.* 44(4):857-61
Lucini D, Cannone V, Malacarne M, Bruno D, Beltrami S et al. 2008. Evidence of autonomic dysregulation in otherwise healthy cancer caregivers: a possible link with health hazard. *Eur. J. Cancer.* 44(16):2437-43
Lucini D, Di FG, Parati G, Pagani M. 2005. Impact of chronic psychosocial stress on autonomic cardiovascular regulation in otherwise healthy subjects. *Hypertension.* 46(5):1201-6
Lucini D, Norbiato G, Clerici M, Pagani M. 2002. Hemodynamic and autonomic adjustments to real life stress conditions in humans. *Hypertension.* 39(1):184-8
Lucini D, Riva S, Pizzinelli P, Pagani M. 2007. Stress management at the worksite: reversal of symptoms profile and cardiovascular dysregulation. *Hypertension.* 49(2):291-7
Luger A, Deuster PA, Kyle SB, Gallucci WT, Montgomery LC et al. 1987. Acute hypothalamic-pituitary-adrenal responses to the stress of treadmill exercise. Physiologic adaptations to physical training. *N. Engl. J. Med.* 316(21):1309-15
Lund I, Lundeberg T, Carleson J, Sonnerfors H, Uhrlin B, Svensson E. 2006. Corticotropin releasing factor in urine--a possible biochemical marker of fibromyalgia. Responses to massage and guided relaxation. *Neurosci. Lett.* 403(1-2):166-71
Lyall M, Peakman M, Wessely S. 2003. A systematic review and critical evaluation of the immunology of chronic fatigue syndrome. *J. Psychosom. Res.* 55(2):79-90
Macedo JA, Hesse J, Turner JD, Meyer J, Hellhammer DH, Muller CP. 2008. Glucocorticoid sensitivity in fibromyalgia patients: decreased expression of corticosteroid receptors and glucocorticoid-induced leucine zipper. *Psychoneuroendocrinology.* 33(6):799-809
MacHale SM, Cavanagh JT, Bennie J, Carroll S, Goodwin GM, Lawrie SM. 1998. Diurnal variation of adrenocortical activity in chronic fatigue syndrome. *Neuropsychobiology.* 38(4):213-7
Macy EM, Hayes TE, Tracy RP. 1997. Variability in the measurement of C-reactive protein in healthy subjects: implications for reference intervals and epidemiological applications. *Clin. Chem.* 43(1):52-8
Madsen T, Christensen JH, Toft E, Schmidt EB. 2007. C-reactive protein is associated with heart rate variability. *Ann. Noninvasive. Electrocardiol.* 12(3):216-22

Maes M, Lin A, Bonaccorso S, van HF, Van GA et al. 1998. Increased 24-hour urinary cortisol excretion in patients with post-traumatic stress disorder and patients with major depression, but not in patients with fibromyalgia. *Acta Psychiatr. Scand.* 98(4):328-35
Maier SF, Watkins LR. 1998. Cytokines for psychologists: implications of bidirectional immune-to-brain communication for understanding behavior, mood, and cognition. *Psychol. Rev.* 105(1):83-107
Majer M, Jones JF, Unger ER, Youngblood LS, Decker MJ et al. 2007. Perception versus polysomnographic assessment of sleep in CFS and non-fatigued control subjects: results from a population-based study. *BMC. Neurol.* 7:40
Male D. 2001. Introduction to the immune system. In *Immunology*, ed. I Roitt, J Brostoff, D Male, 1-13. Mosby.
Maloney EM, Gurbaxani BM, Jones JF, de Souza CL, Pennachin C, Goertzel BN. 2006. Chronic fatigue syndrome and high allostatic load. *Pharmacogenomics.* 7(3):467-73
Malpas SC, Maling TJ. 1990. Heart-rate variability and cardiac autonomic function in diabetes. *Diabetes.* 39(10):1177-81
Malt EA, Olafsson S, Lund A, Ursin H. 2002. Factors explaining variance in perceived pain in women with fibromyalgia. *BMC. Musculoskelet. Disord.* 3:12
Mantel N, Haenszel W. 1959. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* 22(4):719-48
Mantella RC, Butters MA, Amico JA, Mazumdar S, Rollman BL et al. 2008. Salivary cortisol is associated with diagnosis and severity of late-life generalized anxiety disorder. *Psychoneuroendocrinology.* 33(6):773-81
Marsland AL, Gianaros PJ, Prather AA, Jennings JR, Neumann SA, Manuck SB. 2007. Stimulated production of proinflammatory cytokines covaries inversely with heart rate variability. *Psychosom. Med.* 69(8):709-16
Martinez-Lavin M, Hermosillo AG, Mendoza C, Ortiz R, Cajigas JC et al. 1997. Orthostatic sympathetic derangement in subjects with fibromyalgia. *J. Rheumatol.* 24(4):714-8
Masi CM, Hawkley LC, Rickett EM, Cacioppo JT. 2007. Respiratory sinus arrhythmia and diseases of aging: obesity, diabetes mellitus, and hypertension. *Biol. Psychol.* 74(2):212-23
Mayer EA. 2000. The neurobiology of stress and gastrointestinal disease. *Gut.* 47(6):861-9
Mayer EA, Naliboff BD, Chang L, Coutinho SV. 2001. V. Stress and irritable bowel syndrome. *Am. J. Physiol. Gastrointest. Liver Physiol.* 280(4):G519-G524
Mayou R, Farmer A. 2002. ABC of psychological medicine: Functional somatic symptoms and syndromes. *BMJ.* 325(7358):265-8
Mayou R, Kirmayer LJ, Simon G, Kroenke K, Sharpe M. 2005. Somatoform disorders: time for a new approach in DSM-V. *Am. J. Psychiatry* 162(5):847-55
McBeth J, Chiu YH, Silman AJ, Ray D, Morriss R et al. 2005. Hypothalamic-pituitary-adrenal stress axis function and the relationship with chronic widespread pain and its antecedents. *Arthritis Res. Ther.* 7(5):R992-R1000
McBeth J, Silman AJ, Gupta A, Chiu YH, Ray D et al. 2007. Moderation of psychosocial risk factors through dysfunction of the hypothalamic-pituitary-adrenal stress axis in the onset of chronic widespread musculoskeletal pain: findings of a population-based prospective cohort study. *Arthritis Rheum.* 56(1):360-71
McCain GA, Tilbe KS. 1989. Diurnal hormone variation in fibromyalgia syndrome: a comparison with rheumatoid arthritis. *J. Rheumatol. Suppl.* 19:154-7
McDade TW, Hawkley LC, Cacioppo JT. 2006. Psychosocial and behavioral predictors of inflammation in middle-aged and older adults: the Chicago health, aging, and social relations study. *Psychosom. Med.* 68(3):376-81
McEwen BS. 1998a. Protective and damaging effects of stress mediators. *N. Engl. J. Med.* 338(3):171-9
McEwen BS. 1998b. Stress, adaptation, and disease. Allostasis and allostatic load. *Ann. N. Y. Acad. Sci.* 840:33-44.:33-44
McEwen BS. 2007. Physiology and neurobiology of stress and adaptation: central role of the brain. *Physiol. Rev.* 87(3):873-904
McEwen BS, Biron CA, Brunson KW, Bulloch K, Chambers WH et al. 1997. The role of adrenocorticoids as modulators of immune function in health and disease: neural, endocrine and immune interactions. *Brain Res. Brain Res. Rev.* 23(1-2):79-133
McEwen BS, Stellar E. 1993. Stress and the individual. Mechanisms leading to disease. *Arch. Intern. Med.* 153(18):2093-101

McKenzie R, O'Fallon A, Dale J, Demitrack M, Sharma G et al. 1998. Low-dose hydrocortisone for treatment of chronic fatigue syndrome: a randomized controlled trial. *JAMA.* 280(12):1061-6
McLean SA, Williams DA, Harris RE, Kop WJ, Groner KH et al. 2005. Momentary relationship between cortisol secretion and symptoms in patients with fibromyalgia. *Arthritis Rheum.* 52(11):3660-9
McLean SA, Williams DA, Stein PK, Harris RE, Lyden AK et al. 2006. Cerebrospinal fluid corticotropin-releasing factor concentration is associated with pain but not fatigue symptoms in patients with fibromyalgia. *Neuropsychopharmacology.* 31(12):2776-82
Meade MO, Richardson WS. 1997. Selecting and appraising studies for a systematic review. *Ann. Intern. Med.* 127(7):531-7
Meeus M, Nijs J, Van de WN, Toeback L, Truijen S. 2008. Diffuse noxious inhibitory control is delayed in chronic fatigue syndrome: an experimental study. *Pain.* 139(2):439-48
Meewisse ML, Reitsma JB, de Vries GJ, Gersons BP, Olff M. 2007. Cortisol and post-traumatic stress disorder in adults: systematic review and meta-analysis. *Br. J. Psychiatry.* 191:387-92
Mericq MV, Cutler GB. 1998. High fluid intake increases urine free cortisol excretion in normal subjects. *J. Clin. Endocrinol. Metab.* 83(2):682-4
Michaud K, Matheson K, Kelly O, Anisman H. 2008. Impact of stressors in a natural context on release of cortisol in healthy adult humans: a meta-analysis. *Stress.* 11(3):177-97
Milaneschi Y, Corsi AM, Penninx BW, Bandinelli S, Guralnik JM, Ferrucci L. 2009. Interleukin-1 Receptor Antagonist and Incident Depressive Symptoms Over 6 Years in Older Persons: The InCHIANTI Study. *Biol. Psychiatry.* 65(11):973-8
Miller GE, Chen E, Zhou ES. 2007. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. *Psychol. Bull.* 133(1):25-45
Miller GE, Cohen S. 2001. Psychological interventions and the immune system: a meta-analytic review and critique. *Health Psychol.* 20(1):47-63
Mohamed-Ali V, Goodrick S, Rawesh A, Katz DR, Miles JM et al. 1997. Subcutaneous adipose tissue releases interleukin-6, but not tumor necrosis factor-alpha, in vivo. *J. Clin. Endocrinol. Metab.* 82(12):4196-200
Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* 354(9193):1896-900
Moher D, Liberati A, Tetzlaff J, Altman DG. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS. Med.* 6(7):e1000097
Monster TB, Janssen WM, De Jong PE, de Jong-van den Berg LT. 2002. Pharmacy data in epidemiological studies: an easy to obtain and reliable tool. *Pharmacoepidemiol. Drug Saf.* 11(5):379-84
Moorkens G, Berwaerts J, Wynants H, Abs R. 2000. Characterization of pituitary function with emphasis on GH secretion in the chronic fatigue syndrome. *Clin. Endocrinol. (Oxf).* 53(1):99-106
Morriss RK, Robson MJ, Deakin JF. 2002. Neuropsychological performance and noradrenaline function in chronic fatigue syndrome under conditions of high arousal. *Psychopharmacology (Berl).* 163(2):166-73
Moss-Morris R, Spence M. 2006. To "lump" or to "split" the functional somatic syndromes: can infectious and emotional risk factors differentiate between the onset of chronic fatigue syndrome and irritable bowel syndrome? *Psychosom. Med.* 68(3):463-9
Mulder LJM. 1988. Assessment of cardiovascular reactivity by means of spectral analysis. Dissertation, University of Groningen, Groningen, the Netherlands
Muthen LK, Muthen BO. 2004. *Mplus User's Guide,* Los Angeles, CA, Muthen & Muthen
Mutsuura H, Kanbara K, Fukunaga M, Yamamoto K, Ban I et al. 2009. Depression and Anxiety Correlate Differently with Salivary Free Cortisol in the Morning in Patients with Functional Somatic Syndrome. *Appl. Psychophysiol. Biofeedback.* In press
Nater UM, Maloney E, Boneva RS, Gurbaxani BM, Lin JM et al. 2008a. Attenuated morning salivary cortisol concentrations in a population-based study of persons with chronic fatigue syndrome and well controls. *J. Clin. Endocrinol. Metab.* 93(3):703-9
Nater UM, Youngblood LS, Jones JF, Unger ER, Miller AH et al. 2008b. Alterations in diurnal salivary cortisol rhythm in a population-based sample of cases with chronic fatigue syndrome. *Psychosom. Med.* 70(3):298-305

Nault MA, Milne B, Parlow JL. 2002. Effects of the selective H1 and H2 histamine receptor antagonists loratadine and ranitidine on autonomic control of the heart. *Anesthesiology.* 96(2):336-41
Neumann L, Lerner E, Glazer Y, Bolotin A, Shefer A, Buskila D. 2008. A cross-sectional study of the relationship between body mass index and clinical characteristics, tenderness measures, quality of life, and physical functioning in fibromyalgia patients. *Clin. Rheumatol.* 27(12):1543-7
Ng C, Malcolm A, Hansen R, Kellow J. 2007. Feeding and colonic distension provoke altered autonomic responses in irritable bowel syndrome. *Scand. J. Gastroenterol.* 42(4):441-6
Nicolson NA. 2007. Measurement of Cortisol. In *Handbook of Physiological Research Methods in Health Psychology*, ed. LJ Luecken, LC Gallo, 3:37-74. Sage Publications, Inc.
Niewoehner CB, Bantle J. 1998. The adrenal cortex. In *Endocrine Pathophysiology*,61-76. Connecticut, Fence Creek Publishing.
Nimnuan C, Hotopf M, Wessely S. 2001a. Medically unexplained symptoms: an epidemiological study in seven specialities. *J. Psychosom. Res.* 51(1):361-7
Nimnuan C, Rabe-Hesketh S, Wessely S, Hotopf M. 2001b. How many functional somatic syndromes? *J. Psychosom. Res.* 51(4):549-57
Nolan RP, Kamath MV, Floras JS, Stanley J, Pang C et al. 2005. Heart rate variability biofeedback as a behavioral neurocardiac intervention to enhance vagal heart rate control. *Am. Heart J.* 149(6):1137
O'Connor MF, Bower JE, Cho HJ, David CJ, Dimitrov S et al. 2009. To assess, to control, to exclude: Effects of biobehavioral factors on circulating inflammatory markers. *Brain Behav. Immun.* 23(7):887-97
Ohlin B, Nilsson PM, Nilsson JA, Berglund G. 2004. Chronic psychosocial stress predicts long-term cardiovascular morbidity and mortality in middle-aged men. *Eur. Heart J.* 25(10):867-73
Orr WC, Elsenbruch S, Harnish MJ. 2000. Autonomic regulation of cardiac function during sleep in patients with irritable bowel syndrome. *Am. J. Gastroenterol.* 95(10):2865-71
Otte C, Neylan TC, Pole N, Metzler T, Best S et al. 2005. Association between childhood trauma and catecholamine response to psychological stress in police academy recruits. *Biol. Psychiatry.* 57(1):27-32
Ottenweller JE, Sisto SA, McCarty RC, Natelson BH. 2001. Hormonal responses to exercise in chronic fatigue syndrome. *Neuropsychobiology.* 43(1):34-41
Pagani M, Lombardi F, Guzzetti S, Sandrone G, Rimoldi O et al. 1984. Power spectral density of heart rate variability as an index of sympatho-vagal interaction in normal and hypertensive subjects. *J. Hypertens. Suppl.* 2(3):S383-S385
Paine P, Kishor J, Worthen SF, Gregory LJ, Aziz Q. 2009. Exploring relationships for visceral and somatic pain with autonomic control and personality. *Pain.* 144(3):236-44
Paiva ES, Deodhar A, Jones KD, Bennett R. 2002. Impaired growth hormone secretion in fibromyalgia patients: evidence for augmented hypothalamic somatostatin tone. *Arthritis Rheum.* 46(5):1344-50
Papadopoulos A, Ebrecht M, Roberts AD, Poon L, Rohleder N, Cleare AJ. 2009. Glucocorticoid receptor mediated negative feedback in chronic fatigue syndrome using the low dose (0.5 mg) dexamethasone suppression test. *J. Affect. Disord.* 112(1-3):289-94
Pariante CM. 2009. Chronic fatigue syndrome and the immune system: "findings in search of meanings". *Brain Behav. Immun.* 23(3):325-6
Pariante CM, Thomas SA, Lovestone S, Makoff A, Kerwin RW. 2004. Do antidepressants regulate how cortisol affects the brain? *Psychoneuroendocrinology.* 29(4):423-47
Patacchioli FR, Angelucci L, Dellerba G, Monnazzi P, Leri O. 2001. Actual stress, psychopathology and salivary cortisol levels in the irritable bowel syndrome (IBS). *J. Endocrinol. Invest* 24(3):173-7
Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, III et al. 2003. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation.* 107(3):499-511
Pennebaker JW. 1982. *The psychology of physical symptoms,* New York, Springer
Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. 2007. Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Stat. Med.* 26(25):4544-62

Peto R, Pike MC, Armitage P, Breslow NE, Cox DR et al. 1977. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br. J. Cancer.* 35(1):1-39

Petry NM, Barry D, Pietrzak RH, Wagner JA. 2008. Overweight and obesity are associated with psychiatric disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychosom. Med.* 70(3):288-97

Piche ME, Lemieux S, Weisnagel SJ, Corneau L, Nadeau A, Bergeron J. 2005. Relation of high-sensitivity C-reactive protein, interleukin-6, tumor necrosis factor-alpha, and fibrinogen to abdominal adipose tissue, blood pressure, and cholesterol and triglyceride levels in healthy postmenopausal women. *Am. J. Cardiol.* 96(1):92-7

Pinto-Sietsma SJ, Janssen WM, Hillege HL, Navis G, De ZD, De Jong PE. 2000. Urinary albumin excretion is associated with renal functional abnormalities in a nondiabetic population. *J. Am. Soc. Nephrol.* 11(10):1882-8

Pitsavos C, Panagiotakos DB, Papageorgiou C, Tsetsekou E, Soldatos C, Stefanadis C. 2006. Anxiety in relation to inflammation and coagulation markers, among healthy adults: the ATTICA study. *Atherosclerosis.* 185(2):320-6

Pole N, Neylan TC, Otte C, Metzler TJ, Best SR et al. 2007. Associations between childhood trauma and emotion-modulated psychophysiological responses to startling sounds: a study of police cadets. *J. Abnorm. Psychol.* 116(2):352-61

Posserud I, Agerforz P, Ekman R, Bjornsson ES, Abrahamsson H, Simren M. 2004. Altered visceral perceptual and neuroendocrine response in patients with irritable bowel syndrome during mental stress. *Gut.* 53(8):1102-8

Pou KM, Massaro JM, Hoffmann U, Vasan RS, Maurovich-Horvat P et al. 2007. Visceral and subcutaneous adipose tissue volumes are cross-sectionally related to markers of inflammation and oxidative stress: the Framingham Heart Study. *Circulation.* 116(11):1234-41

Prins JB, Bleijenberg G, Bazelmans E, Elving LD, de Boo TM et al. 2001. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet.* 357(9259):841-7

Prins JB, van der Meer JW, Bleijenberg G. 2006. Chronic fatigue syndrome. *Lancet* 367(9507):346-55

Pruessner JC, Wolf OT, Hellhammer DH, Buske-Kirschbaum A, von AK et al. 1997. Free cortisol levels after awakening: a reliable biological marker for the assessment of adrenocortical activity. *Life Sci.* 61(26):2539-49

Racciatti D, Guagnano MT, Vecchiet J, De Remigis PL, Pizzigallo E et al. 2001. Chronic fatigue syndrome: circadian rhythm and hypothalamic-pituitary-adrenal (HPA) axis impairment. *Int. J. Immunopathol. Pharmacol.* 14(1):11-5

Raison CL, Lin JM, Reeves WC. 2009. Association of peripheral inflammatory markers with chronic fatigue in a population-based sample. *Brain Behav. Immun.* 23(3):327-37

Raj SR, Brouillard D, Simpson CS, Hopman WM, Abdollah H. 2000. Dysautonomia among patients with fibromyalgia: a noninvasive assessment. *J. Rheumatol.* 27(11):2660-5

Ranjit N, ez-Roux AV, Shea S, Cushman M, Seeman T et al. 2007. Psychosocial factors and inflammation in the multi-ethnic study of atherosclerosis. *Arch. Intern. Med.* 167(2):174-81

Raven PW, Taylor NF. 1996. Sex differences in the human metabolism of cortisol. *Endocr. Res.* 22(4):751-5

Rechlin T. 1994. The effect of amitriptyline, doxepin, fluvoxamine, and paroxetine treatment on heart rate variability. *J. Clin. Psychopharmacol.* 14(6):392-5

Ridker PM, Rifai N, Stampfer MJ, Hennekens CH. 2000. Plasma concentration of interleukin-6 and the risk of future myocardial infarction among apparently healthy men. *Circulation.* 101(15):1767-72

Riedel W, Layka H, Neeck G. 1998. Secretory pattern of GH, TSH, thyroid hormones, ACTH, cortisol, FSH, and LH in patients with fibromyalgia syndrome following systemic injection of the relevant hypothalamic-releasing hormones. *Z. Rheumatol.* 57 Suppl 2:81-7

Riedel W, Schlapp U, Leck S, Netter P, Neeck G. 2002. Blunted ACTH and cortisol responses to systemic injection of corticotropin-releasing hormone (CRH) in fibromyalgia: role of somatostatin and CRH-binding protein. *Ann. N. Y. Acad. Sci.* 966:483-90

Rief W, Auer C. 2000. Cortisol and somatization. *Biol. Psychol.* 53(1):13-23

Rief W, Barsky AJ. 2005. Psychobiological perspectives on somatoform disorders. *Psychoneuroendocrinology.* 30(10):996-1002

Rief W, Broadbent E. 2007a. Explaining medically unexplained symptoms-models and mechanisms. *Clin. Psychol. Rev.* 27(7):821-41
Rief W, Hiller W, Heuser J. 1997. *SOMS-Das Screening fur Somatoforme Storungen: Manual zum Fragebogen (SOMS-The Screening for Somatoform Symptoms),* Bern, Huber-Verlag.
Rief W, Isaac M. 2007b. Are somatoform disorders 'mental disorders'? A contribution to the current debate. *Curr. Opin. Psychiatry.* 20(2):143-6
Rief W, Nanke A, Emmerich J, Bender A, Zech T. 2004. Causal illness attributions in somatoform disorders: associations with comorbidity and illness behavior. *J. Psychosom. Res.* 57(4):367-71
Rief W, Pilger F, Ihle D, Bosmans E, Egyed B, Maes M. 2001. Immunological differences between patients with major depression and somatization syndrome. *Psychiatry Res.* 105(3):165-74
Rief W, Shaw R, Fichter MM. 1998. Elevated levels of psychophysiological arousal and cortisol in patients with somatization syndrome. *Psychosom. Med.* 60(2):198-203
Ritz T, Dahme B. 2006. Implementation and interpretation of respiratory sinus arrhythmia measures in psychosomatic medicine: practice against better evidence? *Psychosom. Med.* 68(4):617-27
Robbins JM, Kirmayer LJ, Hemami S. 1997. Latent variable models of functional somatic distress. *J Nerv. Ment. Dis.* 185(10):606-15
Robert JJ, Orr WC, Elsenbruch S. 2004. Modulation of sleep quality and autonomic functioning by symptoms of depression in women with irritable bowel syndrome. *Dig. Dis. Sci.* 49(7-8):1250-8
Roberts AD, Charler ML, Papadopoulos A, Wessely S, Chalder T, Cleare AJ. 2009a. Does hypocortisolism predict a poor response to cognitive behavioural therapy in chronic fatigue syndrome? *Psychol. Med.* 39:1-8
Roberts AD, Papadopoulos AS, Wessely S, Chalder T, Cleare AJ. 2009b. Salivary cortisol output before and after cognitive behavioural therapy for chronic fatigue syndrome. *J. Affect. Disord.* 115(1-2):280-6
Roberts AD, Wessely S, Chalder T, Papadopoulos A, Cleare AJ. 2004. Salivary cortisol response to awakening in chronic fatigue syndrome. *Br. J. Psychiatry* 184:136-41
Roberts-Thomson IC, Knight RE, Kennaway DJ, Pannall PR. 1988. Circadian rhythms in patients with abdominal pain syndromes. *Aust. New Zealand J. Med.* 18(4):569-74
Robins LN, Helzer JE, Ratcliff KS, Seyfried W. 1982. Validity of the diagnostic interview schedule, version II: DSM-III diagnoses. *Psychol. Med.* 12(4):855-70
Roelofs K, Spinhoven P. 2007. Trauma and medically unexplained symptoms towards an integration of cognitive and neuro-biological accounts. *Clin. Psychol. Rev.* 27(7):798-820
Rosengren A, Hawken S, Ounpuu S, Sliwa K, Zubaid M et al. 2004. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet.* 364(9438):953-62
Rosenthal R. 1991. Meta-analysis: a review. *Psychosom. Med.* 53(3):247-71
Rosenthal R, DiMatteo MR. 2001. Meta-analysis: recent developments in quantitative methods for literature reviews. *Annu. Rev. Psychol.* 52:59-82.:59-82
Rosmalen JG, Neeleman J, Gans RO, de Jonge P. 2007. The association between neuroticism and self-reported common somatic symptoms in a population cohort. *J. Psychosom. Res.* 62(3):305-11
Rothman KJ. 1976. Causes. *Am. J. Epidemiol.* 104(6):587-92
Rottenberg J. 2007. Cardiac vagal control in depression: a critical analysis. *Biol. Psychol.* 74(2):200-11
Routledge HC, Chowdhary S, Coote JH, Townend JN. 2002. Cardiac vagal response to water ingestion in normal human subjects. *Clin. Sci. (Lond).* 103(2):157-62
Rowbottom D, Keast D, Pervan Z, Goodman C, Bhagat C et al. 1998. The role of glutamine in the aetiology of the chronic fatigue syndrome: A prospective study. *Chronic Fatigue Syndr.* 4(2):3-22
Rozanski A, Blumenthal JA, Kaplan J. 1999. Impact of psychological factors on the pathogenesis of cardiovascular disease and implications for therapy. *Circulation.* 99(16):2192-217
Rubin GJ, Hotopf M, Papadopoulos A, Cleare A. 2005. Salivary cortisol as a predictor of postoperative fatigue. *Psychosom. Med.* 67(3):441-7
Rucker G, Schwarzer G, Carpenter J. 2008. Arcsine test for publication bias in meta-analyses with binary outcomes. *Stat. Med.* 27(5):746-63
Rutledge T, Hogan BE. 2002. A quantitative review of prospective evidence linking psychological factors with hypertension development. *Psychosom. Med.* 64(5):758-66

Ryan SM, Goldberger AL, Pincus SM, Mietus J, Lipsitz LA. 1994. Gender- and age-related differences in heart rate dynamics: are women more complex than men? *J. Am. Coll. Cardiol.* 24(7):1700-7
Sagami Y, Shimada Y, Tayama J, Nomura T, Satake M et al. 2004. Effect of a corticotropin releasing hormone receptor antagonist on colonic sensory and motor function in patients with irritable bowel syndrome. *Gut.* 53(7):958-64
Sajadieh A, Nielsen OW, Rasmussen V, Hein HO, Hansen JF. 2006. C-reactive protein, heart rate variability and prognosis in community subjects with no apparent heart disease. *J. Intern. Med.* 260(4):377-87
Salmon P, Calderbank S. 1996. The relationship of childhood physical and sexual abuse to adult illness behavior. *J. Psychosom. Res.* 40(3):329-36
Sampson M, McGowan J. 2006. Errors in search strategies were identified by type and frequency. *J. Clin. Epidemiol.* 59(10):1057-63
Sandercock G. 2007. Normative values, reliability and sample size estimates in heart rate variability. *Clin. Sci. (Lond).* 113(3):129-30
Sandercock GR, Bromley PD, Brodie DA. 2005a. Effects of exercise on heart rate variability: inferences from meta-analysis. *Med. Sci. Sports Exerc.* 37(3):433-9
Sandercock GR, Bromley PD, Brodie DA. 2005b. The reliability of short-term measurements of heart rate variability. *Int. J. Cardiol.* 103(3):238-47
Sanderman R, Eysenck SBG, Arrindell WA. 1991. Cross-cultural comparisons of personality - the Netherlands and England. *Psychol. Rep.* 69:1091-6
Sanderson S, Tatt ID, Higgins JP. 2007. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int. J. Epidemiol.* 36(3):666-76
Schoepfer AM, Trummler M, Seeholzer P, Seibold-Schmid B, Seibold F. 2008. Discriminating IBD from IBS: comparison of the test performance of fecal markers, blood leukocytes, CRP, and IBD antibodies. *Inflamm. Bowel. Dis.* 14(1):32-9
Schommer NC, Hellhammer DH, Kirschbaum C. 2003. Dissociation between reactivity of the hypothalamus-pituitary-adrenal axis and the sympathetic-adrenal-medullary system to repeated psychosocial stress. *Psychosom. Med.* 65(3):450-60
Schroeder EB, Liao D, Chambless LE, Prineas RJ, Evans GW, Heiss G. 2003. Hypertension, blood pressure, and heart rate variability: the Atherosclerosis Risk in Communities (ARIC) study. *Hypertension.* 42(6):1106-11
Schule C. 2007. Neuroendocrinological mechanisms of actions of antidepressant drugs. *J. Neuroendocrinol.* 19(3):213-26
Schwartz S. 1994. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *Am. J. Public Health.* 84(5):819-24
Scott LV, Burnett F, Medbak S, Dinan TG. 1998a. Naloxone-mediated activation of the hypothalamic-pituitary-adrenal axis in chronic fatigue syndrome. *Psychol. Med.* 28(2):285-93
Scott LV, Dinan TG. 1998d. Urinary free cortisol excretion in chronic fatigue syndrome, major depression and in healthy volunteers. *J. Affect. Disord.* 47(1-3):49-54
Scott LV, Medbak S, Dinan TG. 1998b. Blunted adrenocorticotropin and cortisol responses to corticotropin-releasing hormone stimulation in chronic fatigue syndrome. *Acta Psychiatr. Scand.* 97(6):450-7
Scott LV, Medbak S, Dinan TG. 1998c. The low dose ACTH test in chronic fatigue syndrome and in health. *Clin. Endocrinol. (Oxf).* 48(6):733-7
Scott LV, Medbak S, Dinan TG. 1999a. Desmopressin augments pituitary-adrenal responsivity to corticotropin-releasing hormone in subjects with chronic fatigue syndrome and in healthy volunteers. *Biol. Psychiatry.* 45(11):1447-54
Scott LV, Salahuddin F, Cooney J, Svec F, Dinan TG. 1999b. Differences in adrenal steroid profile in chronic fatigue syndrome, in depression and in health. *J. Affect. Disord.* 54(1-2):129-37
Scott LV, Svec F, Dinan T. 2000. A preliminary study of dehydroepiandrosterone response to low-dose ACTH in chronic fatigue syndrome and in healthy subjects. *Psychiatry Res.* 97(1):21-8
Shapiro S. 1994. Meta-analysis/Shmeta-analysis. *Am J Epidemiol.* 140(9):771-8
Segerstrom SC, Miller GE. 2004. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychol. Bull.* 130(4):601-30
Sharpe M, Bass C. 1992. Pathophysiological mechanisms in somatization. *Int. Rev. Psuchiatry* 4(1):81-97

Sheehan B, Bass C, Briggs R, Jacoby R. 2004. Do general practitioners believe that their older patients physical symptoms are somatized? *J. Psychosom. Res.* 56(3):313-6
Shiri R, Solovieva S, Husgafvel-Pursiainen K, Taimela S, Saarikoski LA et al. 2008. The association between obesity and the prevalence of low back pain in young adults: the Cardiovascular Risk in Young Finns Study. *Am. J. Epidemiol.* 167(9):1110-9
Shufflebotham J, Wetherell MA, Hince D, Hood S, Lightman S et al. 2009. Women with diarrhoea-predominant irritable bowel syndrome show an increased pressure response to 35% carbon dioxide stress challenge. *Stress.* 12(1):30-6
Siegfried N, Muller M, Volmink J, Deeks J, Egger M et al. 2003. Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database Syst. Rev.*(3):CD003362
Simon G, Gater R, Kisely S, Piccinelli M. 1996. Somatic symptoms of distress: an international primary care study. *Psychosom. Med.* 58(5):481-8
Simon GE, Gureje O. 1999. Stability of somatization disorder and somatization symptoms among primary care patients. *Arch. Gen. Psychiatry.* 56(1):90-5
Singer JD, Willett JB. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence,* Oxford University Press.
Singh P, Hawkley LC, McDade TW, Cacioppo JT, Masi CM. 2009. Autonomic tone and C-reactive protein: a prospective population-based study. *Clin. Auton. Res.* 19(6):367-74.
Sipinkova I, Hahn G, Meyer M, Tadlanek M, Hajek J. 1997. Effect of respiration and posture on heart rate variability. *Physiol. Res.* 46(3):173-9
Sloan RP, Shapiro PA, Bagiella E, Boni SM, Paik M et al. 1994. Effect of mental stress throughout the day on cardiac autonomic control. *Biol. Psychol.* 37(2):89-99
Sloan RP, Shapiro PA, DeMeersman RE, Bagiella E, Brondolo EN et al. 2009. The effect of aerobic training and cardiac autonomic regulation in young adults. *Am. J. Public Health.* 99(5):921-8
Smith GD, Ben-Shlomo Y, Beswick A, Yarnell J, Lightman S, Elwood P. 2005. Cortisol, testosterone, and coronary heart disease: prospective evidence from the Caerphilly study. *Circulation.* 19;112(3):332-40
Snijders TAB, Bosker RJ. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling,* Sage Publishers
Sondermeijer HP, van Marle AG, Kamen P, Krum H. 2002. Acute effects of caffeine on heart rate variability. *Am. J. Cardiol.* 90(8):906-7
Speckens AE, Van Hemert AM, Bolk JH, Rooijmans HG, Hengeveld MW. 1996. Unexplained physical symptoms: outcome, utilization of medical care and associated factors. *Psychol. Med.* 26(4):745-52
Spence VA, Kennedy G, Belch JJ, Hill A, Khan F. 2008. Low-grade inflammation and arterial wave reflection in patients with chronic fatigue syndrome. *Clin. Sci. (Lond).* 114(8):561-6
Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV, III et al. 1994. Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA.* 272(22):1749-56
Steffen PR, Smith TB, Larson M, Butler L. Acculturation to Western society as a risk factor for high blood pressure: a meta-analytic review. 2006. *Psychosom. Med.* 68(3):386-97
Stein PK, Domitrovich PP, Ambrose K, Lyden A, Fine M et al. 2004. Sex effects on heart rate variability in fibromyalgia and Gulf War illness. *Arthritis Rheum.* 51(5):700-8
Stein PK, Domitrovich PP, Hui N, Rautaharju P, Gottdiener J. 2005. Sometimes higher heart rate variability is not better heart rate variability: results of graphical and nonlinear analyses. *J Cardiovasc. Electrophysiol.* 16(9):954-9
Sterne JA, Egger M. 2001. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J. Clin. Epidemiol.* 54(10):1046-55
Sterne JA, Gavaghan D, Egger M. 2000. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J. Clin. Epidemiol.* 53(11):1119-29
Stewart JC, Janicki-Deverts D, Muldoon MF, Kamarck TW. 2008. Depressive symptoms moderate the influence of hostility on serum interleukin-6 and C-reactive protein. *Psychosom. Med.* 70(2):197-204
Stone J, Wojcik W, Durrance D, Carson A, Lewis S et al. 2002. What should we say to patients with symptoms unexplained by disease? The "number needed to offend". *BMJ.* 325(7378):1449-50
Streiner DL, Norman GR. 2003. *Health measurement scales: a practical guide to their development and use,* Oxford University Press

Strickland P, Morriss R, Wearden A, Deakin B. 1998. A comparison of salivary cortisol in chronic fatigue syndrome, community depression and healthy controls. *J. Affect. Disord.* 47(1-3):191-4
Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD et al. 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA.* 283(15):2008-12
Sullivan PF, Smith W, Buchwald D. 2002. Latent class analysis of symptoms associated with chronic fatigue syndrome and fibromyalgia. *Psychol. Med.* 32(5):881-8
Sundin EC, Horowitz MJ. Horowitz's Impact of Event Scale evaluation of 20 years of use. 2003. *Psychosom. Med.* 65(5):870-6
Sumathipala A. 2007. What is the evidence for the efficacy of treatments for somatoform disorders? A critical review of previous intervention studies. *Psychosom. Med.* 69(9):889-900
Sutton AJ, Higgins JP. 2008. Recent developments in meta-analysis. *Stat. Med.* 27(5):625-50
Swartz M, Blazer D, Woodbury M, George L, Landerman R. 1986. Somatization disorder in a US southern community: use of a new procedure for analysis of medical classification. *Psychol. Med.* 16(3):595-609
Tak LM, Riese H, De Bock GH, Manoharan A, Kok IC, Rosmalen JG. 2009a. As good as it gets? A meta-analysis and systematic review of methodological quality of heart rate variability studies in functional somatic disorders. *Biol. Psychol.* 82(2):101-10
Tak LM, Rosmalen JGM. 2007. Psychoneuroendocrinology of functional somatic disorders. In *Psychoneuroendocrinology Research Trends*, ed. MT Czberska, 17:463-495. Hauppage, New York, Nova Science Publishers, Inc.
Tak LM, Rosmalen JGM, Bakker SJL. 2009b. Dysfunction of the hypothalamic-pituitary-adrenal-axis and functional somatic symptoms: A longitudinal cohort study in the general population. *Psychoneuroendocrinology.* 34(6):869-77
Tak LM, van Roon AM, Rosmalen JG. 2009c. Is baroreflex sensitivity really pathological in patients with somatization disorder? Comment on: Laederach-Hofmann et al. (2008). *Biol. Psychol.* 82(1):97-8
Tak LM, Cleare AJ, Ormel J, Manoharan A, Kok IC, Wessely S, Rosmalen JG. 2009d. Meta-analysis and meta-regression of HPA axis function in functional somatic disorders. *Submitted.*
Tang YY, Ma Y, Fan Y, Feng H, Wang J et al. 2009. Central and autonomic nervous system interaction is altered by short-term meditation. *Proc. Natl. Acad. Sci. U. S. A.* 106(22):8865-70
Tanriverdi F, Karaca Z, Unluhizarci K, Kelestimur F. 2007. The hypothalamo-pituitary-adrenal axis in chronic fatigue syndrome and fibromyalgia syndrome. *Stress.* 10(1):13-25
Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. 1996. Heart rate variability: standards of measurement, physiological interpretation and clinical use. *Circulation.* 93(5):1043-65
Tayama J, Sagami Y, Shimada Y, Hongo M, Fukudo S. 2007. Effect of alpha-helical CRH on quantitative electroencephalogram in patients with irritable bowel syndrome. *Neurogastroenterol. Motil.* 19(6):471-83
Taylor RL, Machacek D, Singh RJ. 2002. Validation of a high-throughput liquid chromatography-tandem mass spectrometry method for urinary cortisol and cortisone. *Clin. Chem.* 48(9):1511-9
Tepper BJ, Tepper K. 1993. The effects of method variance within measures. *J. Psychol.* 127(3):293-302
Ter Wolbeek M, van Doornen LJ, Coffeng LE, Kavelaars A, Heijnen CJ. 2007. Cortisol and severe fatigue: a longitudinal study in adolescent girls. *Psychoneuroendocrinology.* 32(2):171-82
Thayer JF, Hall M, Sollers JJ, III, Fischer JE. 2006. Alcohol use, urinary cortisol, and heart rate variability in apparently healthy men: Evidence for impaired inhibitory control of the HPA axis in heavy drinkers. *Int. J. Psychophysiol.* 59(3):244-50
Thayer JF, Lane RD. 2007. The role of vagal function in the risk for cardiovascular disease and mortality. *Biol. Psychol.* 74(2):224-42
Thompson JJ. 2003. Autonomic and endocrine activity in irritable bowel syndrome patients with and without depressive symptoms. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 64(4-B):1936
Thompson JJ, Elsenbruch S, Harnish MJ, Orr WC. 2002. Autonomic functioning during REM sleep differentiates IBS symptom subgroups. *Am. J. Gastroenterol.* 97(12):3147-53
Thompson SG. 1994. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ.* 19;309(6965):1351-5

Tillisch K, Mayer EA, Labus JS, Stains J, Chang L, Naliboff BD. 2005. Sex-specific alterations in autonomic function among patients with irritable bowel syndrome. *Gut.* 54(10):1396-401
Torpy DJ, Bachmann AW, Gartside M, Grice JE, Harris JM et al. 2004. Association between chronic fatigue syndrome and the corticosteroid-binding globulin gene ALA SER224 polymorphism. *Endocr. Res.* 30(3):417-29
Torpy DJ, Papanicolaou DA, Lotsikas AJ, Wilder RL, Chrousos GP, Pillemer SR. 2000. Responses of the sympathetic nervous system and the hypothalamic-pituitary-adrenal axis to interleukin-6: a pilot study in fibromyalgia. *Arthritis Rheum.* 43(4):872-80
Tousignant-Laflamme Y, Goffaux P, Bourgault P, Marchand S. 2006. Different autonomic responses to experimental pain in IBS patients and healthy controls. *J. Clin. Gastroenterol.* 40(9):814-20
Tsuji H, Larson MG, Venditti FJ, Jr., Manders ES, Evans JC et al. 1996. Impact of reduced heart rate variability on risk for cardiac events. The Framingham Heart Study. *Circulation.* 94(11):2850-5
Turk CS, Gatz M, Kato K, Pedersen NL. 2008. Physical health 25 years later: the predictive ability of neuroticism. *Health Psychol.* 27(3):369-78
Turnbull AV, Rivier CL. 1999. Regulation of the hypothalamic-pituitary-adrenal axis by cytokines: actions and mechanisms of action. *Physiol. Rev.* 79(1):1-71
Uceyler N, Hauser W, Sommer C. 2008. A systematic review on the effectiveness of treatment with antidepressants in fibromyalgia syndrome. *Arthritis Rheum.* 59(9):1279-98
Ukkola O, Gagnon J, Rankinen T, Thompson PA, Hong Y et al. 2001. Age, body mass index, race and other determinants of steroid hormone variability: the HERITAGE Family Study. *Eur. J. Endocrinol.* 145(1):1-9
Ursin H. 1997. Sensitization, somatization, and subjective health complaints. *Int. J. Behav. Med.* 4(2):105-16
Van Cauter E, Leproult R, Kupfer DJ. 1996. Effects of gender and age on the levels and circadian rhythmicity of plasma cortisol. *J. Clin. Endocrinol. Metab.* 81(7):2468-73
Van Den Eede F. 2007. Hypothalamic-pituitary-adrenal axis function in major depressive disorder and in chronic fatigue syndrome. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 68(4-B):2241
Van Den Eede F, Moorkens G, Hulstijn W, Van Houdenhove B, Cosyns P et al. 2008. Combined dexamethasone/corticotropin-releasing factor test in chronic fatigue syndrome. *Psychol. Med.* 38(7):963-73
Van Denderen JC, Boersma JW, Zeinstra P, Hollander AP, van Neerbos BR. 1992. Physiological effects of exhaustive physical exercise in primary fibromyalgia syndrome (PFS): is PFS a disorder of neuroendocrine reactivity? *Scand. J. Rheumatol.* 21(1):35-7
Van der Werf SP, Prins JB, Vercoulen JH, van der Meer JW, Bleijenberg G. 2000. Identifying physical activity patterns in chronic fatigue syndrome using actigraphic assessment. *J. Psychosom. Res.* 49(5):373-9
Van Eekelen AP, Houtveen JH, Kerkhof GA. 2004a. Circadian variation in base rate measures of cardiac autonomic activity. *Eur. J. Appl. Physiol.* 93(1-2):39-46
Van Houdenhove B, Egle U, Luyten P. 2005. The role of life stress in fibromyalgia. *Curr. Rheumatol. Rep.* 7(5):365-70
Van Houdenhove B, Egle UT. 2004. Fibromyalgia: a stress disorder? Piecing the biopsychosocial puzzle together. *Psychother. Psychosom.* 73(5):267-75
Van Leeuwen N, Kumsta R, Entringer S, de Kloet ER, Zitman FG et al. 2009. Functional mineralocorticoid receptor (MR) gene variation influences the cortisol awakening response after dexamethasone. *Psychoneuroendocrinology.* In press
Van Melle JP, de Jonge P, Spijkerman TA, Tijssen JG, Ormel J et al. 2004. Prognostic association of depression following myocardial infarction with mortality and cardiovascular events: a meta-analysis. *Psychosom. Med.* 66(6):814-22
Van Orshoven NP, Andriesse GI, van Schelven LJ, Smout AJ, Akkermans LM, Oey PL. 2006. Subtle involvement of the parasympathetic nervous system in patients with irritable bowel syndrome. *Clin. Auton. Res.* 16(1):33-9
Van Ravenswaaij-Arts CM, Kollee LA, Hopman JC, Stoelinga GB, van Geijn HP. 1993. Heart rate variability. *Ann. Intern. Med.* 118(6):436-47
Van Rensburg SJ, Potocnik FC, Kiss T, Hugo F, van ZP et al. 2001. Serum concentrations of some metals and steroids in patients with chronic fatigue syndrome with reference to neurological and cognitive abnormalities. *Brain Res. Bull.* 55(2):319-25

Van Wijk CM, Kolk AM. 1997. Sex differences in physical symptoms: the contribution of symptom perception theory. *Soc. Sci. Med.* 45(2):231-46
Vercoulen JH, Bazelmans E, Swanink CM, Fennis JF, Galama JM et al. 1997. Physical activity in chronic fatigue syndrome: assessment and its role in fatigue. *J. Psychiatr. Res.* 31(6):661-73
Videlock EJ, Adeyemo M, Licudine A, Hirano M, Ohning G, Mayer M et al. 2009. Childhood Trauma Is Associated With Hypothalamic-Pituitary-Adrenal Axis Responsiveness in Irritable Bowel Syndrome. *Gastroenterology.* In press.
Virtanen R, Jula A, Huikuri H, Kuusela T, Helenius H et al. 2004. Increased pulse pressure is associated with reduced baroreflex sensitivity. *J. Hum. Hypertens.* 18(4):247-52
Visser J, Graffelman W, Blauw B, Haspels I, Lentjes E et al. 2001. LPS-induced IL-10 production in whole blood cultures from chronic fatigue syndrome patients is increased but supersensitive to inhibition by dexamethasone. *J. Neuroimmunol.* 119(2):343-9
Visser M, Bouter LM, McQuillan GM, Wener MH, Harris TB. 1999. Elevated C-reactive protein levels in overweight and obese adults. *JAMA.* 282(22):2131-5
Vollmer-Conna U, Aslakson E, White PD. 2006. An empirical delineation of the heterogeneity of chronic unexplained fatigue in women. *Pharmacogenomics.* 7(3):355-64
Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. 2007. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology.* 18(6):800-4
Vreeburg SA, Hoogendijk WJ, van Pelt J, DeRijk RH, Verhagen JC et al. 2009a. Major depressive disorder and hypothalamic-pituitary-adrenal axis activity: results from a large cohort study. *Arch. Gen. Psychiatry.* 66(6):617-26
Vreeburg SA, Kruijtzer BP, van Pelt J, van Dyck R, Derijk RH et al. 2009b. Associations between sociodemographic, sampling and health factors and various salivary cortisol indicators in a large sample without psychopathology. *Psychoneuroendocrinology.* 34(8):1109-20
Waal de MW, Arnold IA, Eekhof JA, Van Hemert AM. 2004. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. *Br. J. Psychiatry.* 184:470-6
Wallace DJ. 2006. Is there a role for cytokine based therapies in fibromyalgia. *Curr. Pharm. Des.* 12(1):17-22
Walter SA, ardal-Eriksson E, Thorell LH, Bodemar G, Hallbook O. 2006. Pre-experimental stress in patients with irritable bowel syndrome: high cortisol values already before symptom provocation with rectal distensions. *Neurogastroenterol. Motil.* 18(12):1069-77
Wang H, Moser M, Schiltenwolf M, Buchner M. 2008. Circulating cytokine levels compared to pain in patients with fibromyalgia -- a prospective longitudinal study over 6 months. *J. Rheumatol.* 35(7):1366-70
Waring WS, Chui M, Japp A, Nicol EF, Ford MJ. 2004. Autonomic cardiovascular responses are impaired in women with irritable bowel syndrome. *J. Clin. Gastroenterol.* 38(8):658-63
Weissbecker I, Floyd A, Dedert E, Salmon P, Sephton S. 2006. Childhood trauma and diurnal cortisol disruption in fibromyalgia syndrome. *Psychoneuroendocrinology.* 31(3):312-24
Wessely S, Nimnuan C, Sharpe M. 1999. Functional somatic syndromes: one or many? *Lancet.* 354(9182):936-9
Wessely S, White PD. 2004. There is only one functional somatic syndrome. *Br. J. Psychiatry* 185:95-6
Wiebe N, Vandermeer B, Platt RW, Klassen TP, Moher D, Barrowman NJ. 2006. A systematic review identifies a lack of standardization in methods for handling missing variance data. *J. Clin. Epidemiol.* 59(4):342-53
Wieseler-Frank J, Maier SF, Watkins LR. 2005. Immune-to-brain communication dynamically modulates pain: physiological and pathological consequences. *Brain Behav. Immun.* 19(2):104-11
Wijeratne C, Brodaty H, Hickie I. 2003. The neglect of somatoform disorders by old age psychiatry: some explanations and suggestions for future research. *Int. J. Geriatr. Psychiatry.* 18(9):812-9
Wilcox RR. 2002. Summarizing Data: Detecting Outliers. In *Applying contemporary statistical techniques*, 3:55-92. San Diego, Academic Press.
Wileman L, May C, Chew-Graham CA. 2002. Medically unexplained symptoms and the problem of power in the primary care consultation: a qualitative study. *Fam. Pract.* 19(2):178-82
Wing JK, Babor T, Brugha T, Burke J, Cooper JE et al. 1990. SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch. Gen. Psychiatry.* 47(6):589-93

Wingenfeld K, Heim C, Schmidt I, Wagner D, Meinlschmidt G, Hellhammer DH. 2008. HPA axis reactivity and lymphocyte glucocorticoid sensitivity in fibromyalgia syndrome and chronic pelvic pain. *Psychosom. Med.* 70(1):65-72
Wingenfeld K, Wagner D, Schmidt I, Meinlschmidt G, Hellhammer DH, Heim C. 2007. The low-dose dexamethasone suppression test in fibromyalgia. *J. Psychosom. Res.* 62(1):85-91
Wise TN. 2009. Diagnostic Criteria for Psychosomatic Research Are Necessary for DSM V. *Psychother. Psychosom.* 78(6):330-2
Wolfe F, Hawley DJ. 1998. Psychosocial factors and the fibromyalgia syndrome. *Z. Rheumatol.* 57 Suppl 2:88-91
Wood B, Wessely S, Papadopoulos A, Poon L, Checkley S. 1998. Salivary cortisol profiles in chronic fatigue syndrome. *Neuropsychobiology.* 37(1):1-4
Wood PB. 2005. Neuroimaging in functional somatic syndromes. *Int. Rev. Neurobiol.* 67:119-63
Wulsin LR, Singal BM. Do depressive symptoms increase the risk for the onset of coronary disease? A systematic quantitative review. 2003. *Psychosom. Med.* 65(2):201-10
Wust S, Federenko I, Hellhammer DH, Kirschbaum C. 2000. Genetic factors, perceived chronic stress, and the free cortisol response to awakening. *Psychoneuroendocrinology.* 25(7):707-20
Yamamoto Y, LaManca JJ, Natelson BH. 2003. A measure of heart rate variability is sensitive to orthostatic challenge in women with chronic fatigue syndrome. *Exp. Biol. Med. (Maywood).* 228(2):167-74
Yataco A, Talo H, Rowe P, Kass DA, Berger RD, Calkins H. 1997. Comparison of heart rate variability in patients with chronic fatigue syndrome and controls. *Clin. Auton. Res.* 7(6):293-7
Yatham LN, Morehouse RL, Chisholm BT, Haase DA, MacDonald DD, Marrie TJ. 1995. Neuroendocrine assessment of serotonin (5-HT) function in chronic fatigue syndrome. *Can. J. Psychiatry.* 40(2):93-6
Yotsukura M, Koide Y, Fujii K, Tomono Y, Katayama A et al. 1998. Heart rate variability during the first month of smoking cessation. *Am. Heart J.* 135(6 Pt 1):1004-9
Young AH, Sharpe M, Clements A, Dowling B, Hawton KE, Cowen PJ. 1998. Basal activity of the hypothalamic-pituitary-adrenal axis in patients with the chronic fatigue syndrome (neurasthenia). *Biol. Psychiatry.* 43(3):236-7
Zarkovic M, Pavlovic M, Pokrajac-Simeunovic A, Ciric J, Beleslin B et al. 2003. Disorder of adrenal gland function in chronic fatigue syndrome. *Srp. Arh. Celok. Lek.* 131(9-10):370-4

SUMMARY

Somatization is defined as the experience of somatic symptoms not explained by conventional underlying organic pathology. These symptoms may occur in isolation (functional somatic symptoms, FSS) or cluster in syndromes of related complaints (functional somatic disorders, FSD), such as chronic fatigue syndrome (CFS), fibromyalgia (FM), or irritable bowel syndrome (IBS). Psychological, biological, and social factors are involved in the etiology of somatization. Psychosocial stress is widely accepted as a psychological factor playing a role in the etiology of somatization, but underlying mechanisms linking psychosocial stress to somatization are largely unknown. One of the potential biological factors is dysfunction of stress responsive systems, namely, the autonomic nervous system (ANS), the hypothalamic-pituitary-adrenal (HPA) axis, and the immune system. A reason to investigate a potential etiological link between the stress responsive systems and somatization emerges from the potential of those systems to increase symptom experience and perception. Although the theory of dysfunction of stress responsive systems in somatization is intuitively attractive, as it could explain why psychosocial stress may result in somatization, research findings of the last twenty years have been conflicting. It is yet unknown whether dysfunction of stress responsive systems plays a causal role in the etiology of somatization. Mixed previous findings may be due to the fact that many studies in this research area have been performed in small samples and have a large variety in designs. Meta-analysis is a particularly useful technique to quantitatively review these findings (general principles of meta-analysis and its specific merits and pitfalls for observational, psychosomatic research are outlined in *Chapter 2*).

In *Chapter 1*, we hypothesized that dysfunctions in stress responsive systems, as characterized by hypofunction of the HPA axis, altered ANS function in the direction of decreased cardiac vagal activity, and increased immune system activation, are risk factors in the etiology of somatization. We investigated these hypotheses with meta-analyses on FSD and population-based cohort studies on FSS.

In the meta-analysis on dysfunction of the ANS and FSD *(Chapter 3),* we included studies that measured cardiac vagal activity using heart rate variability (HRV) in patients with FSD and healthy controls. The systematic review of overall methodological quality of those HRV studies in FSD (i.e., CFS, FM, and IBS) urged a need for methodological improvements in this area. Meta-analysis of 14 studies confirmed our hypothesis of lower cardiac vagal activity in FSD patients compared to controls. However, the reliability of this summary estimate was significantly limited by unexplained heterogeneity and potential publication bias. Although overall evidence suggests reduced parasympathetic activity as a generic factor, available evidence was not adequate to firmly reject or accept a role of ANS dysfunction in FSD. Therefore, quality criteria and recommendations to improve future research on HRV in FSD were provided in this chapter.

In the meta-analysis on dysfunction of the HPA axis and FSD *(Chapter 4)*, we hypothesized that FSD would be characterized by hypocortisolism. Meta-analysis of 85 studies revealed that although basal cortisol levels were generally lower in FSD subjects compared to controls, this association did not reach statistical significance. However, when we calculated the association with the three FSD separately, statistically significant hypocortisolism was observed in CFS subjects compared to controls, but not in FM or IBS. Female gender, medication use, and presence of co-morbid depressive disorder were moderators of the association between hypocortisolism and FSD. When all potential moderators were entered into a meta-regression analysis, only type of FSD and female gender were significant independent predictors of hypocortisolism. This meta-analysis highlighted the relevance of further identifying hypocortisolemic subgroups among and within FSD.

As the meta-analyses were only based on cross-sectional case-control studies, which often failed to control for confounders, we aimed to translate the findings on FSD in our meta-analysis to a longitudinal study on the experience of FSS in the general population. The second part of this thesis was therefore based on data derived from a population-based cohort of 1094 adults. In this cohort, we were able to control for several potential confounders of the association between dysfunction of stress responsive systems and somatization, such as gender, age, body mass index, smoking, depressive disorder, and physical inactivity. Furthermore, as two waves of data were available, we could also evaluate whether stress responsive system dysfunction, if present, also preceded onset of FSS, or was merely a concomitant or consequence. Participants in this study completed the somatization section of the Composite International Diagnostic Interview (CIDI) surveying the presence of 43 FSS. As a primary continuous outcome measure, we used the total number of FSS. Secondly, as previously suggested clustering of cardiopulmonary, musculoskeletal, gastrointestinal, and general FSS could be replicated in *Chapter 5*, we also used FSS in these bodily clusters as an outcome measure. The musculoskeletal cluster includes FM symptoms and the gastrointestinal cluster includes IBS symptoms. Although fatigue is not surveyed in the somatization section of the CIDI, the general cluster includes secondary symptoms as seen in CFS.

We aimed to investigate the hypothesis that cardiac vagal activity is related to somatization in the general population in *Chapter 6*. Cardiac vagal activity was assessed by spectral analysis of HRV. We found that lower cardiac vagal activity was associated with the experience of FSS in younger adults. This association was similar for FSS in general, musculoskeletal, and gastrointestinal bodily clusters. In older adults, unexpectedly, higher cardiac vagal activity was associated with the experience of FSS. Longitudinal analysis demonstrated a similar pattern. We concluded that, in line with our meta-analysis based on studies which predominantly included patients younger than 50 years, decreased cardiac vagal

activity is associated with a higher number of FSS in younger adults in the general population. The unexpected association between increased cardiac vagal activity and FSS in older adults needs further exploration.

In the same population-based cohort, no association was found between HPA axis function, as indexed by 24-hour urinary cortisol excretion, and the total number of FSS. No association was found for 24-hour urinary free cortisol and general, musculoskeletal, or gastrointestinal bodily clusters of FSS. Furthermore, 24-hour urinary cortisol excretion did not predict new-onset FSS. We conclude that the study, which is presented in *Chapter 7*, does not provide evidence for an association between altered HPA axis function and FSS in the general population, which contrasts with our hypotheses as generated by the meta-analysis on HPA axis function and FSD.

Whether dysfunction of the immune system, a third bodily stress responsive system, is associated with somatization was investigated in *Chapter 8*. Low-grade immune system activation was indexed with high sensitive C-reactive protein (hs-CRP). Similar to our results on 24-hour urinary cortisol, hs-CRP was not associated with the total number of FSS. When examining different bodily clusters of FSS, however, hs-CRP was significantly associated with musculoskeletal FSS, such as back pain, joint pain, pain in extremities, and muscle weakness. Elevated hs-CRP also preceded new-onset of these symptoms. This differential association may be explained by the fact that musculoskeletal FSS are specifically related to sickness behavior, mainly to the hyperalgesia and weakness components. However, the value of hs-CRP as an adequate biomarker for sickness behavior in the general population remains to be established.

In aggregate, we demonstrated that alterations in stress responsive systems might contribute to somatization in subgroups of patients. One of the premises of this thesis was that psychosocial stress, which is widely accepted to contribute to the etiology of somatization, affects functioning of the stress responsive systems. Accordingly, *Chapter 9* aims to study the effects of cumulative psychosocial stress on stress responsive systems, by examining whether self-reported adverse life events during the lifespan are associated with current activity of ANS, HPA axis, or immune system. Life events were measured by a slightly adapted version of the List of Threatening Experiences, surveying events such as loss of close relative, marital difficulties, or a major financial crisis. We observed a negative association between the lifetime score of adverse life events and cardiac vagal activity, but not with 24-hour urinary cortisol or hs-CRP. We concluded that our data provide no support for an association between psychosocial stress and dysfunction of stress responsive systems in the general population, with the exception of a modest association between adverse life-events and ANS function. This study contradicts the generally accepted idea that psychosocial stress influences stress responsive systems.

We conclude that dysfunction of stress responsive systems is not a necessary factor for somatization in general. However, lowered cardiac vagal activity, hypocortisolism, and possibly low-grade immune activation, may be relevant in the etiology in subgroups of patients with FSD. Such subgroups need to be better identified in order to target and improve treatment strategies based on normalization of function of the stress responsive systems.

LEKENSAMENVATTING

Iedereen ervaart wel eens lichamelijke klachten. Soms is hiervoor een duidelijke oorzaak, zoals een beschadiging of een ontsteking, maar vaak is er geen duidelijke medische verklaring te vinden. Deze onverklaarde lichamelijke klachten kunnen op zichzelf voorkomen; dan noemen we ze met een vakterm 'functioneel somatische symptomen'. Deze klachten kunnen ook samen in clusters voorkomen; dan noemen we het 'functioneel somatische syndromen'. Voorbeelden hiervan zijn het chronisch vermoeidheid syndroom (CVS), fibromyalgie (FM) en het prikkelbare darm syndroom (PDS). Patiënten met CVS hebben last van moeheid, maar ook keelpijn en duizeligheid kunnen voorkomen. Symptomen die bij FM voorkomen zijn pijn in gewrichten en spieren, maar ook moeheid kan een rol spelen. Bij het PDS kunnen patiënten last hebben van buikpijn, een veranderde stoelgang en een opgeblazen gevoel.

Het ervaren van lichamelijke symptomen die niet verklaard kunnen worden door een bekende medische oorzaak wordt ook wel somatisatie genoemd. Somatisatie is een groot probleem in de gezondheidszorg. Terwijl artsen niet goed weten hoe ze deze klachten het best kunnen behandelen, voelen patiënten met deze klachten zich vaak niet begrepen. Daarom is het van groot belang het ontstaan van somatisatie beter te begrijpen. Hoewel de meeste onderzoekers vermoeden dat het doormaken van stressvolle gebeurtenissen of moeilijkheden ('psychosociale stress') een risicofactor is voor het ontstaan van somatisatie, is het niet goed bekend welke mechanismen hieraan ten grondslag liggen. Een mogelijk mechanisme zou verstoring kunnen zijn van de systemen in het menselijk lichaam die op stress reageren (stress responsieve systemen). Drie belangrijke stress responsieve system zijn het autonoom zenuwstelsel (ANZ), de hypothalamus-hypofyse-bijnier as (HHB as) en het immuun systeem. Het ANZ, dat vrijwel met alle organen in verbinding staat via zenuwen, bestaat uit een parasympathisch 'rust' deel en een sympathisch 'stress' deel. Psychosociale stress kan leiden tot een verlaging van de parasympathische activiteit en een verhoging van de sympathische activiteit. De HHB as heeft als eindproduct het stress-hormoon cortisol. Terwijl acute stress (bijv. overvallen worden op straat) meestal een verhoging van het cortisol niveau geeft, lijkt chronische stress (bijv. zorg voor een zieke naaste of werkloosheid) juist te resulteren in een verlaging van het cortisol niveau ('hypocortisolisme'). Het immuun systeem, ten slotte, beschermt niet alleen het lichaam tegen ziekteverwekkers, maar kan ook worden geactiveerd door psychosociale stress. Een subtiele verhoging van het ontstekingseiwit C-reactive protein (hs-CRP) is een maat voor laaggradige activatie van het immuun systeem.

In *Hoofdstuk 1* veronderstelden we dat verstoringen in de drie stress responsieve systemen, gekarakteriseerd door een veranderd functioneren van de HHB as (laag cortisol), het ANZ (verminderde parasympathische activiteit) en het immuun systeem (laaggradige activatie), risicofactoren zijn voor somatisatie. Een reden om geïnteresseerd te zijn in deze verstoringen is dat ze mogelijk kunnen leiden

tot het ervaren van lichamelijke symptomen. De onderzoeksbevindingen van de afgelopen twintig jaar spreken elkaar echter tegen. Het is daardoor nog niet goed bekend of verstoringen in deze stress responsieve systemen nu werkelijk een oorzakelijke rol spelen in somatisatie. De tegenstrijdigheid van voorgaande onderzoeksbevindingen zou veroorzaakt kunnen worden door het feit dat veel studies in dit onderzoeksgebied in kleine groepen patiënten verricht werden. Ook waren ze vaak verschillend in opzet, bijvoorbeeld wat betreft het werven van deelnemers of het meten van verstorende factoren. Meta-analyse is een nuttige statistische techniek om deze bevindingen toch samen te vatten (in *Hoofdstuk 2* bespreken we de algemene principes van meta-analyse, en de specifieke meerwaarde en valkuilen van het gebruik van deze techniek in psychosomatisch onderzoek).

In de meta-analyse die onderzoekt of er veranderingen zijn in het functioneren van het ANZ bij mensen met functioneel somatiche syndromen (*Hoofdstuk 3*), vonden we allereerst dat de kwaliteit van deze studies in veel gevallen voor verbetering vatbaar was. De meta-analyse liet zien dat er inderdaad sprake was van lagere parasympathische activiteit in mensen met functioneel somatische syndromen. Dit verband was hetzelfde voor CVS, FM en PDS. Echter, de studies waren onderling te verschillend om zeker te zijn dat lagere parasympathische activiteit een rol speelt in deze syndromen. Verder waren er ook aanwijzingen dat studies die dit resultaat niet vonden mogelijk wel verricht maar niet gepubliceerd zijn; het beeld dat we hebben kan hierdoor vertekend zijn. Dus hoewel verlaagde parasympathische activiteit inderdaad een rol lijkt te spelen in functioneel somatische syndromen, is het beschikbare bewijs niet voldoende om definitieve conclusies te trekken. In dit hoofdstuk doen we daarom ook aanbevelingen om in de toekomst beter onderzoek op dit gebied te kunnen verrichten.

In de meta-analyse die onderzoekt of veranderingen zijn in het functioneren van de HHB as bij mensen met functioneel somatische syndromen (*Hoofdstuk 4)* vonden we dat het cortisol niveau inderdaad lager was in patiënten met een functioneel somatisch syndroom vergeleken met gezonde personen. Echter, dit verschil was niet statistisch significant. Wanneer we naar de drie functionele somatische syndromen apart keken, vonden we dat er een statistisch significant lager cortisol niveau was in patiënten met CVS, maar niet in patiënten met FM of PDS. Ook vonden we dat een lager cortisol niveau vooral bij vrouwen met functioneel somatische syndromen een rol lijkt te spelen. Uit deze meta-analyse blijkt dat niet iedereen met somatisatie een lager cortisol niveau heeft, maar dat het van belang is om te onderzoeken bij welke subgroepen van patiënten een verlaagd cortisol niveau wel een rol speelt.

Beide meta-analyses waren gebaseerd op studies die maar op één punt in de tijd hebben gemeten. Bovendien hielden deze studies dikwijls geen rekening met de rol van factoren die het onderzoeken van de relatie tussen stress responsieve

systemen en somatisatie kunnen verstoren. In het tweede deel van het proefschrift hebben we daarom gekeken of de bevindingen uit de meta-analyses ook konden worden vertaald naar het ervaren van functioneel somatische symptomen in de algemene bevolking.

We gebruikten hiervoor gegevens afkomstig van een groep van 1094 volwassen mannen en vrouwen. In deze studie werd op twee tijdsmomenten gemeten, waardoor we konden bepalen of veranderingen in stress responsieve systemen voorafgingen aan de ontwikkeling van functioneel somatische symptomen. Ook hebben we rekening gehouden met de invloed van verschillende verstorende factoren, zoals leeftijd, geslacht, body mass index, roken, alcoholgebruik, de aanwezigheid van een depressieve stoornis en fysieke inactiviteit. We keken zowel naar de totale hoeveelheid functioneel somatische symptomen die deelnemers in deze studie rapporteerden, als ook naar verschillende clusters van symptomen (zoals bescheven in *Hoofdstuk 5*). Hoewel vermoeidheid niet was gemeten is in deze studie, hebben de symptomen in het algemene cluster, zoals hoofdpijn en duizeligheid, het meest weg van CVS. Het cluster met symptomen van het bewegingsapparaat lijkt op FM, terwijl de symptomen in het maag-darm cluster lijken op die van patiënten met PDS.

In *Hoofdstuk 6* onderzochten we of er een verband was tussen lagere parasympathische activiteit en het ervaren van functioneel somatische symptomen. We vonden dat dit inderdaad het geval was voor jongvolwassenen. Dit verband met lagere parasympathische activiteit was gelijk voor functioneel somatische symptomen in het algemene-, bewegingsapparaat- en maag-darmcluster. Deze bevindingen repliceren het verband van lagere parasympathische activiteit met somatisatie dat we eerder al in de meta-analyse gevonden hadden. Echter, onverwachts vonden we bij oudere volwassenen een verband tussen verhoogde parasympathische activiteit en de hoeveelheid functioneel somatische symptomen. Deze onverwachte bevinding vereist verder onderzoek voordat er conclusies aan verbonden kunnen worden, maar suggereert dat het van belang is de rol van leeftijd in acht te nemen wanneer men het verband tussen het ANZ en somatisatie wil onderzoeken.

In dezelfde bevolkingsstudie vonden we geen verband tussen het functioneren van de HHB as, gemeten via de uitscheiding van het stress-hormoon cortisol in de 24-uurs urine, en de hoeveelheid functioneel somatische symptomen. Ook was er geen verband tussen 24-uurs urine cortisol en functioneel somatische symptomen in het algemene-, bewegingsapparaat- of maag-darmcluster. We concludeerden dus dat deze studie (beschreven in *Hoofdstuk 7*) geen bewijs levert voor een verband tussen veranderingen in HHB as functie en functioneel somatische symptomen. Deze bevinding strookt niet met wat we op basis van de meta-analyse over HHB as functie in patiënten met functioneel somatische syndromen zouden verwachten.

Of activatie van het immuun systeem, het derde stress responsieve systeem van het menselijk lichaam, in verband staat met somatisatie werd onderzocht in *Hoofdstuk 8.* Laaggradige immuun activatie werd bepaald door het meten van het ontstekingseiwit hs-CRP in het bloed. Net als bij het 24-uurs urine cortisol was er geen verband tussen hs-CRP en het totale aantal gerapporteerde functioneel somatische symptomen. Echter, wanneer we naar de verschillende clusters van symptomen keken, vonden we dat hs-CRP wel symptomen zoals gewrichtspijn, spierzwakheid en pijn in armen of benen in het bewegingsapparaat cluster voorspelde. Deze bevinding zou verklaard kunnen worden door het fenomeen 'sickness behavior'. Activatie van het immuun systeem kan leiden tot 'sickness behavior', gekenmerkt door verhoogde pijngevoeligheid, malaise en verhoogde aandacht voor het eigen lichaam: verschijnselen die het meest passen bij symptomen in het bewegingsapparaat-cluster. Het meeste onderzoek naar 'sickness behavior' is echter verricht in dieren en ernstig zieke mensen. Het is dus nog niet goed bekend of hs-CRP een geschikte maat is voor 'sickness behavior' in relatief gezonde mensen in de algemene bevolking.

Alles bij elkaar genomen vonden we dat veranderingen in stress responsieve systemen zouden kunnen bijdragen aan somatisatie in subgroepen patiënten. Eén van de aannames in dit proefschrift was dat psychosociale stress, waarvan velen vermoeden dat het een rol speelt bij somatisatie, invloed heeft op het functioneren van de stress responsieve systemen. Daarom onderzochten we in *Hoofdstuk 9* of dit ook in onze data het geval was. Als een maat voor psychosociale stress vroegen we de deelnemers naar de hoeveelheid negatieve levensgebeurtenissen in hun leven, zoals dood van een naaste, problemen met de levenspartner of het doormaken van een financiële crisis. Zoals verwacht vonden we dat hoe meer negatieve levensgebeurtenissen iemand had doorgemaakt, hoe lager zijn of haar parasympathische activiteit. Er was echter geen relatie met 24-uurs urine cortisol of hs-CRP. Dus, met uitzondering van een verandering in het ANZ, konden onze data het bestaan van een relatie tussen psychosociale stress en stress responsieve systemen niet bevestigen. Aanvullend onderzoek, waarbij grote groepen mensen langere tijd gevolgd worden, zal moeten uitwijzen in hoeverre het idee dat psychosociale stress leidt tot veranderingen in stress responsieve systemen kan standhouden.

We concluderen dat verstoringen in stress responsieve systemen niet altijd hoeven voor te komen bij somatisatie. Een verlaagde parasympathische activiteit, een verlaagd cortisol niveau, en misschien laaggradige immuunactivatie zouden wel van belang kunnen zijn bij het ontstaan van functioneel somatische symptomen en syndromen bij subgroepen patiënten. Zulke subgroepen moeten beter herkend worden, omdat behandeling hierop mogelijk aangepast kan worden.

DANKWOORD

Nog onwetend van eventuele beren op de weg, schreef ik ooit moedig op een groot blauw schrift '*Lineke, 8 jaar, schrijfster*'. Ik vind het daarom toch wel bijzonder dat er nu een proefschrift van mijn hand ligt. Vooral omdat ik eigenlijk nooit een promotie-traject had overwogen voor ik in aanraking kwam met onderzoek naar somatisatie – een even wijdverspreid als fascinerend probleem in de geneeskunde.

Eén van de belangrijkste redenen te gaan promoveren was het idee dat ik veel zou kunnen leren van mijn co-promotor, dr. Rosmalen. Dit was inderdaad het geval. Judith, ik bewonder je durf zaken groots aan te pakken – en standaard te streven naar net iets beter. Je bent uniek in het vinden van creatieve oplossingen. Jouw onderzoekslust is onuitputtelijk; hopelijk ben ik er mee besmet. *Gebaar voor gebaar reikt / ze haar lichaam over aan haar dromen / en telkens wordt ze meer / dan ze al was en telkens reikt ze weer* ('Ballerina', Herman de Coninck).

Mijn promotoren prof. dr. Ormel en prof. dr. Slaets wil ik bedanken voor het kritisch volgen en bevragen van mijn ideeën en plannen. Joris en Hans, jullie geruststellende woorden tijdens voortgangsbesprekingen, namelijk dat er in een vroeg stadium al genoeg lag voor een proefschrift, was tegen dovemansoren gericht. Het maakte wel dat ik vrij kon nadenken wat ik echt belangrijk vond te onderzoeken.

Ik vond het een meerwaarde om te kunnen samenwerken met mensen uit verschillende disciplines. Alle co-auteurs wil ik daarom bedanken voor hun bijdragen aan de papers in dit proefschrift. Dat jullie commentaar nog lang in mijn hoofd moge blijven resoneren bij toekomstig denk- en schrijfwerk.

Ook ben ik de leescommissie, prof. dr. van Doornen, prof. dr. Spinhoven en prof. dr. Gans erkentelijk voor het inhoudelijk beoordelen van het manuscript.

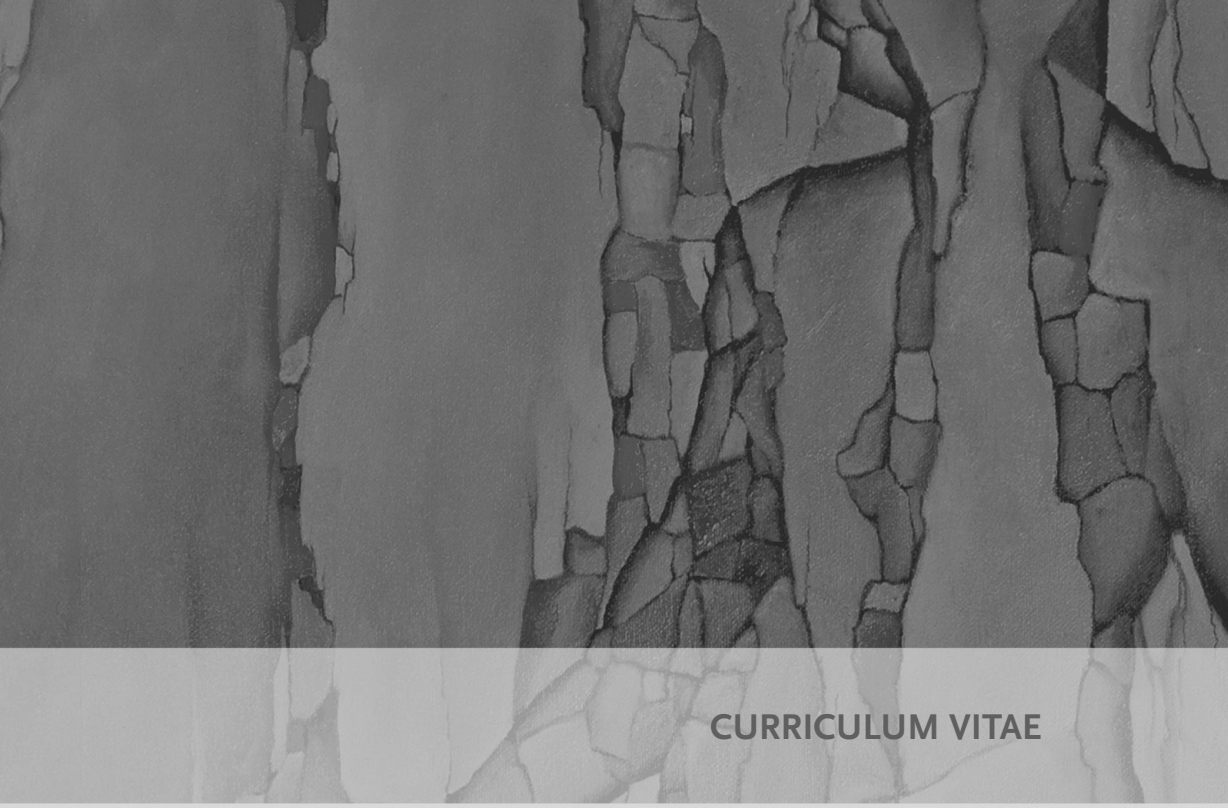
I would like to thank all my colleagues at the Institute of Psychiatry in London, especially prof. dr. Wessely and dr. Cleare, for their help during my six months' visit. The enthusiastic, skillful atmosphere at the Institute of Psychiatry is comparable to the one in Groningen – however, I sadly experienced that your fire alarm is significantly more sensitive.

Helaas heb ik meerdere malen lopen verkondigen dat ik niet van onwaarschijnlijke dankwoorden hou waarin iedereen even bijzonder is. Daarmee heb ik mijn eigen glazen ingegooid, want wat schrijf ik dan, als al mijn Groningse collega's waarlijk stuk voor stuk leuke, eigenzinnige mensen zijn? Ik wil jullie graag warm bedanken voor de goede sfeer op de onderzoeksafdeling Psychiatrie.

Marjolein en Raoul, vanzelfsprekend zijn jullie paranimfen. *Desnieteenstaande, julle is baie wonderlik.*

Een proefschrift is een geheel ander boek dan ik op mijn achtste in gedachten had. Daarvoor wil ik mijn ouders ten slotte bedanken. Bedankt voor de vrijheid open te staan voor het onverwachte. Bedankt voor het relativeren van eventuele beren op de weg. Het is mooi opgegroeid te zijn met het idee dat veel mogelijk is.

Lineke



CURRICULUM VITAE

Lineke Tak was born on April 30th, 1983 in Apeldoorn and was brought up in Loenen, The Netherlands. She studied Medicine at the University Medical Center Utrecht. She did internships at the dermatology department at the Gu-Rankawa Hospital, Medunsa University in Pretoria, South-Africa, at the Rintveld Clinic for Eating Disorders, Altrecht in Zeist, and at an specialized Addiction Care clinic, IrisZorg in Arnhem. She accomplished her main internship at the Medical Psychiatry Unit of the Mesos Medical Center in Utrecht, working on the interface between somatic medicine and psychiatry. While studying medicine, she also conducted a research project on biopsychosocial mechanisms underlying somatization at the Department of Psychiatry, University Medical Center Groningen (UMCG), University of Groningen under supervision of Dr. Judith Rosmalen. She graduated cum laude in 2007. From 2008 till 2010, Lineke did a PhD on dysfunction of stress responsive systems and somatization at the Interdisciplinary Center for Psychiatric Epidemiology (UMCG), University of Groningen under supervision of Dr. Judith Rosmalen, Prof. Dr. Hans Ormel, and Prof. Dr. Joris Slaets. In 2008, she won a Young Investigator Award of the International Society of Psychoneuroendocrinology. In 2009, she worked for six months at the Department of Psychological Medicine at the Institute of Psychiatry, King's College, London, collaborating with Dr. Anthony Cleare and Prof. Dr. Simon Wessely. Currently, Lineke Tak is living in Deventer and will start her specialist training in Psychiatry at Dimence in 2010.

LIST OF PUBLICATIONS

Ockenburg van SL, **Tak LM**, Gans ROB, De Jonge P, Rosmalen JGM. Effects of cumulative stress on stress responsive systems: Results from a population-based cohort study. *Submitted.*

Rosmalen JGM, **Tak LM**, Jonge de P. Empirical foundations for the diagnosis of somatization: implications for DSM-V. *Submitted.*

Tak LM, Cleare AJ, Manoharan A, Kok IC, Ormel J, Rosmalen JGM. Meta-analysis and meta-regression of hypothalamic-pituitary-adrenal axis activity in functional somatic disorders. *Submitted.*

Tak LM, Meijer A, Manoharan A, Rosmalen JGM. More than the sum of its parts: Meta-analysis and its potential to discover sources of heterogeneity in psychosomatic medicine. *Psychosom Med, in press.*

Tak LM, Rosmalen JGM. Dysfunction of stress responsive systems as a risk factor for functional somatic disorders. *J Psychosom Res, in press.*

Tak LM, Janssens KAM, Rosmalen JGM. Age-specific associations between cardiac vagal activity and functional somatic symptoms: A population-based study. *Psychother Psychosom, in press.*

Kingma EM, **Tak LM**, Huisman M, Rosmalen JGM. Intelligence is negatively associated with the number of functional somatic symptoms. *J Epidemiol Community Health 2009;63(11):900-5*

Tak LM, Bakker SJL, Rosmalen JGM. Association between high sensitive C-reactive protein and functional somatic symptoms: a longitudinal population study. *Brain Behav Immun 2009;23(7):1014-9.*

Tak LM, Riese H, de Bock GH, Manoharan A, Kok IC, Rosmalen JGM. As good as it gets? A meta-analysis and systematic review of methodological quality of heart rate variability studies in functional somatic disorders. *Biol Psychol 2009;82(2):101-10.*

Tak LM, Van Roon AM, Rosmalen JGM. Is baroreflex sensitivity really pathological in patients with somatization disorder? Comment on: Laederach-Hofmann et al. *Biol Psychol 2009;82(1):97-8.*

Tak LM, Bakker SJL, Rosmalen JGM. Dysfunction of the hypothalamic-pituitary-adrenal axis and functional somatic symptoms: A longitudinal cohort study in the general population. *Psychoneuroendocrinology 2009;34(6):869-77.*

Kingma EM, **Tak LM**, Rosmalen JGM. Intelligence: A vulnerability factor for functional somatic symptoms? *In: Columbus AM (Ed.), Advances in Psychology Research, Nova Science Publishers, Inc., Hauppauge, New York, 2009, pp. 165-77.*

Tak LM. Vetgemest in de psychiatrie. Co-ervaring - Welke ervaring bepaalt de beroepskeuze van de arts in spe? *Onder redactie van M. Meijer. Van Zuiden Communications, Alphen aan de Rijn, 2008, pp 81-2.*

Tak LM. Commentary: No effect of pharmacological treatment in depressive disorder after myocardial infarction. [Referaat: Geen effect van medicamenteuze behandeling op depressie na een mycoard infarct]. *Ned Tijdschr Geneeskd Stud 2008;(11)1:6-7.*

Tak LM, Rosmalen JGM. Psychoneuroendocrinology of functional somatic disorders. *In: Czberska MT (Ed.), Psychoneuroendocrinology Research Trends. Nova Science Publishers, Inc., Hauppauge, New York, 2007, pp. 463-95.*

SHARE

This thesis is published within the Graduate School for Health Research SHARE (embedded in the University Medical Center Groningen / University of Groningen). More information regarding the institute and its research can be obtained from our internetsite: www.rug.nl/share.

Recent dissertations:

Al Hadithy AFY (2009) Pharmacogenetics of antipsychotic-induced parkinsonism and tardive dyskinesia; a focus on African-Caribbeans and Slovanic Caucasians
Supervisors: prof dr JRBJ Brouwers, prof dr AJM Loonen
Co-supervisors: dr B Wilffert, dr R Bruggeman

Hodselmans AP (2009) Psychophysical capacity in non-specific low back pain
Supervisors: prof dr JHB Geertzen, prof dr PU Dijkstra
Co-supervisor: dr CP van der Schans

Andela AM (2009) Frailty in the clinical practice of nursing care
Supervisors: prof dr R Sanderman, prof dr JPJ Slaets
Co-supervisor: dr A Dijkstra

Henselmans I (2009) Psychological well-being and perceived control after a breast cancer diagnosis
Supervisors: prof dr AV Ranchor, prof dr R Sanderman
Co-supervisor: dr J de Vries

Oud MJT (2009) Zorg van de huisarts voor mensen met ernstige psychische stoornissen
Supervisor: prof dr B Meyboom-de Jong
Co-supervisors: dr J Schuling, dr CJ Slooff

Doormaal JE van (2009) Medication errors and adverse drug events in hospitalised patients; methodological issues and computerised intervention
Supervisor: prof dr FM Haaijer-Ruskamp
Co-supervisors: dr PGM Mol, dr JGW Kosterink, dr PLA van den Bemt

Chang CMS (2009) Ageing with joy; the effect of a physical activity programme on the well-being of older people; a study conducted in five homes for the elderly in Paramaribo
Supervisors: prof dr JR van Horn, prof dr JW Groothoff, prof dr MA Vrede
Co-supervisor: dr M Stevens

Peters-Veluthamaningal C (2009) Corticosteroid injections for the treatment of hand and wrist disorders in general practice
Supervisor: prof dr B Meyboom-de Jong
Co-supervisor: dr JC Winters

Gordijn S (2009) On perinatal pathology; aspects of the perinatal autopsy, placental pathology and classification of perinatal mortality
Supervisors: prof dr JP Holm, prof dr TY Khong
Co-supervisors: dr JJHM Erwich, dr A Timmer

Krokavcová M (2009) Perceived health status in multiple sclerosis patients
Supervisor: prof dr JW Groothoff
Co-supervisors: dr JP van Dijk, dr I Nagyová, dr Z Gdovinová, dr LJ Middel

Beinema MJ (2009) The optimalization of coumarin anticoagulant therapy: pharmacogenetics and computer asisted dose finding
Supervisors: prof dr JRBJ Brouwers, prof dr J van der Meer
Co-supervisor: dr B Wilffert

Boersma C (2009) Health economics of cardiovascular & renal disease prevention
Supervisors: prof dr MJ Postma, prof dr LTW de Jong-van den Berg,
prof dr PE de Jong
Co-supervisor: dr T Gansevoort

El-Sayed Hussein El-Baz N (2009) Effect of clinical pathway implementation and patients' characteristics on outcomes of coronary artery bypass graft surgery
Supervisor: prof dr SA Reijneveld
Co-supervisors: dr LJ Middel, dr JP van Dijk, dr PW Boonstra

Buitenhuis J (2009) The course of whiplash; its psychological determinants and consequences for work disability
Supervisors: prof dr JW Groothoff, prof dr PJ de Jong
Co-supervisor: dr JPC Jaspers

Soer R (2009) Functional capacity evaluation; measurement qualities and normative values
Supervisors: prof dr JHB Geertzen, prof dr JW Groothoff
Co-supervisors: dr MF Reneman, dr CP van der Schans

Knibbe M (2009) 'Not a matter of choice'. Ethical perspectives on decision making about living parental liver donation
Supervisors: prof dr MA Verkerk, prof dr MJH Slooff
Co-supervisor: dr ELM Maeckelberghe

Santvoort MM van (2009) Disability in Europe; policy, social participation and subjective well-being
Supervisor: prof dr WJA van den Heuvel
Co-supervisors: dr JP van Dijk, dr LJ Middel

Stewart RE (2009) A multilevel perspective of patients and general practitioners
Supervisors: prof dr B Meyboom-de Jong, prof dr TAB Snijders, prof dr FM Haaijer-Ruskamp

Vroom F (2009) Drug use in pregnancy; exploring the field of DMARDs in pregnancy
Supervisors: prof dr LTW de Jong-van den Berg, prof dr JRBJ Brouwers, prof dr MAFJ van de Laar
Co-supervisor: dr HEK de Walle

Jong J de (2009) The GALM effect study; changes in physical activity, health and fitness of sedentary and underactive older adults aged 55-65
Supervisor: prof dr EJA Scherder
Co-supervisors: dr KAPM Lemmink, dr M Stevens

Berendsen, A (2009) Samenwerking tussen huisarts en specialist
Supervisor: prof dr B Meyboom-de Jong
Co-supervisor: dr J Schuling